KU Leuven
Biomedical Sciences Group
Faculty of Medicine
Department of Microbiology and Immunology
Laboratory of Virology and Chemotherapy

IDENTIFICATION AND VALIDATION OF DRUG-TARGET INTERACTIONS USING CRISPR/CAS-MEDIATED GENOME EDITING

Jasper Edgar NEGGERS

Jury:

Promoter:	**Prof. Dr. Dirk Daelemans**	*KU Leuven*
Chair:	**Prof. Dr. Johan Neyts**	*KU Leuven*
Secretary	**Prof. Dr. Graciela Andrei**	*KU Leuven*
Jury members:	**Prof. Dr. Michael White**	*University of Texas - Southwestern Medical Center*
	Prof. Dr. Sebastian Nijman	*University of Oxford*
	Prof. Dr. Ludwig Missiaen	*KU Leuven*
	Prof. Dr. Peter Hoet	*KU Leuven*

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Biomedical Sciences

Leuven, April 2018

Dankwoord

Met het schrijven van dit dankwoord krijg ik het gevoel dat ik een levensfase vol met indrukwekkende ervaringen afsluit. Dit hoofdstuk begon, naar mijn gevoel, in principe in de zomer van 2008, toen ik in een redelijk ruim studentenkot op de bovenste verdieping van de Abdij Keizersberg in Leuven introk. Over de jaren heen hebben erg veel mensen mij op verschillende manieren geholpen. Maar aangezien de eerste jaren van dit bijna 10-jarig avontuur al enige tijd geleden zijn, zou ik hier vooral graag mijn collega's, vrienden en begeleiders die betrokken waren bij de tweede helft van mijn Belgische ontdekkingsreis, en bij uitstek vooral bij het tot stand komen van dit doctoraat, willen bedanken.

Eerst en vooral wil ik mijn promotor Dirk bedanken. Dirk heeft mij vanaf het begin van mijn labrotaties en masterthesis altijd gesteund en heeft mijn wetenschappelijke carrière tot de dag van vandaag uitstekend begeleid. Meer specifiek hebben zijn visie op de wetenschap en zijn bereidheid en openheid om de uitdagendste en nieuwste laboratoriumtechnieken en biologische inzichten toe te passen mij gevormd tot de jonge en ambitieuze wetenschapper van vandaag. Ook zou ik graag, zoals wellicht elke student onder Dirk in de afgelopen 10 jaar heeft benadrukt, zijn vrijwel onuitputtelijk enthousiasme voor de wetenschap en zijn groepje collega's benadrukken. Dit enthousiasme is van ongekend niveau en is uitermate aanstekelijk. Daarenboven zou ik Dirk ook graag willen bedanken voor het vertrouwen, de flexibiliteit en de vrijheid die hij mij heeft gegeven om mijn doctoraatsonderzoek uit te voeren. Hij heeft mij bovendien ook aangemoedigd en ondersteund om op congres te gaan. Terugkijkend is een betere promotor en begeleider voor de afgelopen jaren eigenlijk onmogelijk in te beelden.

Uiteraard zou ik ook mijn juryleden willen bedanken. Prof. Dr. Graciela Andrei, Prof. Dr. Ludwig Missiaen, Prof. Dr. Sebastian Nijman, Prof. Dr. Michael White en Prof. Dr. Peter Hoet hebben met aandacht dit manuscript nagekeken en voorzien van het nodige commentaar.

Ten derde wil ik graag mijn eerste wetenschappelijke begeleider, latere collega doctorandus en uiteindelijk oud-collega Eline bedanken. Zij heeft mijn, soms wat overdreven, hyper-enthousiasme in het prille begin van mijn labrotaties uitermate goed weten op te vangen. Zonder haar begeleiding en enthousiasme was ik wellicht ook niet bij Dirk terecht gekomen om mijn masterthesis, en later ook dit doctoraat, te starten. Daarnaast heeft zij tijdens mijn masterthesis en de begin jaren van mijn doctoraat een luisterend oor geboden en kon ik bij haar terecht voor kleine vraagjes, filosofische babbels en uiteraard ook voor wetenschappelijke discussies en advies.

Ook zou ik graag mijn tweede begeleider en latere collega Thomas willen bedanken. Allereerst heeft Thomas mij tijdens mijn masterthesis uitstekend begeleid. Tevens zijn zijn wetenschappelijke kennis en nuchtere wereldvisie altijd indrukwekkend en inspirerend geweest. Het was dan ook erg aangenaam om mijn experimentele obstakels en wetenschappelijke vraagstukken met Thomas te bediscussiëren. Zoals eenieder wel weet, konden wij hier soms, uit enthousiasme en misschien ook een beetje uit koppigheid, wel erg ver in opgaan. Verder heeft Thomas, samen met Eline, mij tijdens de moeilijkere momenten van mijn doctoraat weten op te beuren en te ondersteunen. Zijn wetenschappelijk enthousiasme is overigens bijna net zo groot en aanstekelijk als dat van Dirk!

Graag zou ik ook Lotte, de, in mijn ogen, laboratorium manager van het Dirk Daelemans lab, willen bedanken. Lotte is een gestructureerde collega en heeft mij in het begin (samen met Eline) in het lab wegwijs gemaakt. Daarna was zij altijd bereid om mij met mijn praktische vragen en problemen in het lab te helpen. Ook heeft zij haar welgekende en gestructureerde laboratoriumkwaliteiten op cruciale momenten ingezet om mijn onderzoek en experimenten mee vooruit te helpen. Zonder Lotte zou dit doctoraat duidelijk een stukje stroever tot stand zijn gekomen.

Vervolgens wil ik ook Maarten bedanken. Maarten is vanaf het begin van mijn masterthesis en doctoraat altijd een trouwe en fijne collega geweest die soms, samen met Thomas, erg enthousiast over voetbal kan praten. Gelukkig kunnen wij rekenen op zijn hoogstaande humor om van elk gesprek of elke gebeurtenis een niet te vergeten moment te maken (ook mede dankzij zijn foto's). Maarten is daarnaast ook een uitstekende wetenschapper en zijn praktisch inzicht en vernuft om experimenten simpel en toch krachtig en relevant te houden, hebben mij meerdere malen op het juiste been gezet. Men kan ook op Maarten rekenen om elk feestje, etentje, uitje en sociale aangelegenheid in goede banen te leiden.

Vanaf het begin trof Els mij als een rustige, wijze en ervaren wetenschapper en collega die op de achtergrond de verschillende projecten in het lab ondersteunt en begeleidt. Zij was altijd direct beschikbaar om mijn vragen en problemen rond de confocale microscoop op te lossen. Bovendien heeft zij mij ook geholpen met het uitvoeren van een aantal experimenten.

Tevens wil ik Bob, Astrid en Kristien graag bedanken. Alle drie zijn uitstekende collega's die ons laboratorium halverwege mijn doctoraat kwamen versterken. Wie Bob niet goed kent, kan misschien door zijn, naar mijn bescheiden mening, "bad-boy" uitstraling afgeschrikt worden. Niets is echter meer verraderlijk dan iemands uitstraling verkeerd te interpreteren. Bob is namelijk een vrolijke en aangename collega wiens expertise en interesse altijd beschikbaar staan om de mensen te helpen. Astrid heeft ons laboratorium inmiddels verlaten, maar tijdens haar tijd bij ons heeft ze de sfeer en de praktische gang van zaken in het lab een stevige boost weten te geven. Haar motivatie en interesse waren bovendien van ongekend niveau en haar behulpzaamheid heeft mij, en anderen, meerdere malen goed vooruitgeholpen. Kristien kwam ons lab versterken na een klein crisis momentje binnen V&C. Al snel bleek dat Kristien een erg bekwame en behulpzame collega is met een hoge motivatie en wetenschappelijke interesse. Haar hulp en ondersteuning worden dan ook sterk geapprecieerd.

Zonder mijn mededoctoraatstudenten Bert en Shari was mijn doctoraat een stukje saaier en moeizamer verlopen. Bert is een kalme, kritische en slimme collega die altijd beschikbaar is om wetenschappelijke vragen en kwesties te bespreken. Op momenten kan hij ook snoeihard met zijn droge humor uit een onverwachte hoek komen. Shari heeft het lab inmiddels verlaten, maar haar relatief korte tijd met ons was erg aangenaam. Haar betrokkenheid, precisie, motivatie en wetenschappelijke belangstelling waren bovendien een goede toevoeging aan het DD-groepje. Ten slotte beslaat de nieuwste uitbreiding van ons lab een jonge lichting vrouwelijke collega's, Liesbeth, Nathalie en Maxine, en een nieuwe Nederlandse collega, Jim. Alhoewel deze collega's pas recent het DD-groepje hebben versterkt, maken hun veelzijdigheid, behulpzaamheid, motivatie en wetenschappelijke interesse het DD-groepje in ieder geval een nog fijner team om in te werken.

Uiteraard zijn er ook veel wetenschappelijke collega's buiten het DD-groepje die mij ontzettend geholpen hebben tijdens mijn doctoraat. Allereerst zou ik graag de collega's van de groep van Prof. Dr. Christophe Pannecouque willen bedanken. Aan het begin van mijn wetenschappelijke carrière heb ik mijn bureau en laboruimte met Kris, Kristien, Cindy en Leen gedeeld. Zij hebben mij toen vriendelijk ontvangen, ondersteund en onder andere begeleid in het P3-lab. Daarnaast was Christophe zelf meerdere malen betrokken bij onze wetenschappelijke discussies. Vervolgens zou ik ook graag de (oud)-collega's van het 6de en 7de van het oude torengebouw in de onderzoeksgroepen van Prof. Dr. Annemie van Damme, prof. Dr. Kristel van Laethem, Prof. Dr. Jelle Matthijnssen en prof. Dr. Marc Van Ranst willen bedanken. Meer specifiek zou ik graag Tim bedanken. Tim heeft mij als beginneling wegwijs gemaakt in de wonderlijke wereld van R en R-studio en heeft mij een glimp van de mogelijkheden en wonderen binnen het grote vakgebied van de bioinformatica laten zien. Ook wil ik Lowie, Liana, Lies, Soraya, Nádia en Mónica bedanken voor hun vriendschap en ondersteuning aan het begin van mijn doctoraat. Ook Cathy, Inge, Dominique, Chantal en Christiane van het Rega secretariaat verdienen een stevige merci voor hun sublieme en snelle afwerking van al mijn verschillende orders, logistieke vragen en verzendingen.

Dan wil ik graag Sarah bedanken voor haar technische ondersteuning en flexibiliteit bij het gebruik van de MiSeq. In dezelfde adem wil ik ook Wim en Céline bedanken voor hun service en ondersteuning bij de Genomics Core. Verder bedank ik Prof. Dr. Arnout Voet en zijn trouwe postdoc Hiroki Noguchi voor hun hulp met de verfijning en analyse van de kristalstructuur van NAMPT. Tevens bedank ik ook Michel Delforge en Nicolas Kint voor hun hulp met de proteasoom activiteit's assay. Uiteraard wil ik ook Karyopharm Therapeutics bedanken voor de wetenschappelijke ondersteuning en het uitlenen van hun compounds. Zonder de gastvrijheid van de groep van Leentje was ik ook niet ver gekomen aan het einde van mijn doctoraat en dus zou ik ook graag Leentje, Kirsten, Niels, Lies, Bianca en Nathalie willen bedanken.

Buiten mijn doctoraat en de KU Leuven zijn er natuurlijk ook een heleboel mensen die mij buiten het werk ondersteund hebben en betrokken waren bij mijn sociale ontwikkeling. Allereerst zou ik graag mijn trouwe vrienden van de middelbareschooltijd, Thijs, Martin, Ingo, Bram, Tijl en Erwin, willen bedanken. Verder wil ik ook graag de vrienden van het "CampZone team", Koen, Chris, Michel, Sander, Thomas v.d. B., Thomas v.d. V., Mei-Li en Josephine bedanken. Uiteraard verdient mijn familie een stevig bedankje voor alle steun en motivatie die ik ontvangen heb. Meer specifiek wil ik graag mijn broer Reinout bedanken voor zijn interesse en beschikbaarheid om te babbelen. Mijn vader en moeder wil ik ook erg graag bedanken voor hun flexibiliteit, bewondering, vertrouwen, raad, begrip, motivatie en ondersteuning. Ik ben zeer blij dat ik de vrijheid en mogelijkheid heb gekregen om in België te gaan studeren en dit doctoraat te verwezenlijken. In moeilijkere tijden was ik thuis ook altijd welkom en waren jullie bereid om mij te helpen met wat dan ook. Top!

And last, but certainly not least, I would like to thank Katerina for her tremendous support and love. You have always been a very bright light in my life over the last few years. I remember fondly how we got to know each other by running all around Leuven. I have always admired your strong willpower and determination to pursue your goals and ideas and I know it has not been easy for you to go to the US all alone and that I have taken some time to finish this PhD. However, this thesis manuscript is the final step to finish this epic and experience-rich chapter of my life and very soon I will be able to join you in Boston. I love you very much and I will see you very soon!

Many thanks everyone for your support and friendship!

Jasper "Unclejack" Neggers

Table of contents

Abbreviations

ADME	Absorption, distribution, metabolism and excretion
AGO2	Argonaute 2
AID	Activation-induced cytidine deaminase
AML	Acute myeloid leukemia
ANOVA	Analysis of variance
AsCpf1	*Acidaminococcus* species Cpf1
ASH	American society of hematology
ATCC	American type culture collection
BFP	Blue fluorescent protein
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
Cas	CRISPR-associated
CD	Circular dichroism
Chr	Chromosome
cm	Centimetre
CMap	Connectivity map
CMC	Chemistry, manufacturing and control
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CRISPRres	CRISPR-induced resistance selection in essential genes
CRM1	Chromosomal Region maintenance 1
crRNA	CRISPR RNA
DAmP	Decreased abundance by mRNA perturbation
DAPI	4',6-diamidino-2-phenylindole
DARTS	Drug affinity responsive target stability
dCas9	Dead Cas9
Del	Deletion
DGCR8	DiGeorge syndrome chromosomal [or critical] region 8
DIC	Differential interference contrast
DLBCL	Diffuse large B-cell lymphoma
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSB	Double strand break
EC50	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EMA	European Medicines Agency
EMS	Ethylmethane sulfate
ENU	*N*-ethyl-*N*-nitrosourea
ERCC3	Excision repair cross-complementation group 3
FBS	Fetal bovine serum
FC	Fold change
FDA	Food and Drug Administration

FG	Phenlyalanine-glycine
GDP	Guanosine diphosphate
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
HDR	Homology-directed repair
HDx	Hydrogen-deuterium exchange
HEAT	Huntingtin, Elongation Factor 3, Protein Phosphatase 2A and Target of Rapamycin
HeLa	Henrietta Lacks
HIP	Haploinsufficieny profiling
HIV	Human immunodeficiency virus
HOP	Homozygous deletion profiling
HTRF	Homogeneous time resolved fluorescence
IMDM	Iscove's modified Dulbecco's medium
IN	Integrase
IND	Investigational new drug
Indel	Insertion-deletion
Ins	Insertion
Isp	Ispinesib
KEGG	Kyoto encyclopedia of genes and genomes
KIF11	Kinesin family member 11
KO	Knockout
KPNB1	Karyopherin subunit beta 1
KPT	Karyopharm Therapeutics
LDL-R	Low-density lipoprotein receptor
LMB	Leptomycin B
MCL	Mantle cell lymphoma
MDR	Multi-drug resistance
MDS	Myelodysplastic syndrome
mL	Millilitre
mM	Millimolar
MM	Multiple myeloma
MNV	Multi-nucleotide variation
MOE	Molecular operating environment
MOI	Multiplicity of infection
mRNA	Messenger RNA
MSP	Multicopy suppression profiling
NA	Nicotinic acid
NAD	Nicotinamide adenine dinucleotide
NAMPT	Nicotinamide phosphoribosyltransferase
NAPRT1	Nicotinate phosphoribosyltransferase
NCI	National Cancer Institute
NES	Nuclear export signal
NHEJ	Non-homologous end joining
NLS	Nuclear localization signal
nM	Nanomolar
NPC	Nuclear Pore Complex

n.s.	Not significant
NUC	Nuclease
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PAK4	p21-associated kinase 4
PAM	Protospacer adjacent motif
PARP	Poly (ADP-ribose) polymerase
PASS	Prediction of activity spectra for substances
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PD	Pharmacodynamics
PDB	Protein data bank
PFA	Paraformaldehyde
PFS	Protospacer flanking sequence
pH	Potential of hydrogen
PI	Propidium iodide
PK	Pharmacokinetics
PKI	Protein kinase A inhibitor
PLATO	Parallel analysis of translated ORFs
PMBCL	Primary mediastinal B-cell lymphoma
PR	Protease
PSMB5	Proteasome subunit beta type-5
Q-VD-Oph	Quinoline-Val-Asp-difluorophenoxymethyl ketone
RanBP	Ran binding protein
RanGAP	Ran GTPase activating protein
REC	Recognition
RIPA	Radio-immuno precipitation assay
RISC	RNA-induced silencing complex
RLC	RISC-loading complex
RLU	Relative light unit
RNA	Ribonucleic acid
RNAi	RNA interference
RPM	Reads per million
RPMI	Roswell Park Memorial Institute
RPS3a	Ribosomal protein S3a
RT	Reverse transcriptase
s.d.	Standard deviation
SDS	Sodium dodecyl sulfate
SEA	Similarity ensemble approach
SEC-TID	Size-exclusion chromatography for target identification
sgRNA	Single guide RNA
shRNA	Short hairpin RNA
SINE	Selective inhibitors of nuclear transport
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
snRNA	Small nuclear RNA
SNV	Single nucleotide variation

SpCas9	*Streptococcus pyogenes* Cas9
ssDNA	Single-stranded DNA
SV40	Simian Virus 40
TALEN	Transcription activator like effector nuclease
T-ALL	T-cell acute lympoblastic leukemia
TFO	Triplex forming oligonucleotides
TICC	Target identification by chromatographic co-elution
TPP	Thermal proteome profiling
tracrRNA	Tracer RNA
TRBP	Transactivating RNA-binding protein
Trip	Triptolide
US	United States
UTR	Untranslated region
UV	Ultraviolet
VSV-G	Vesicular stomatitis virus G
WT	Wild-type
XPO1	Exportin-1
XPO5	Exportin-5
YFP	Yellow fluorescent protein
ZFN	Zinc-finger nuclease
β-NMN	β-nicotinamide mononucleotide
μM	Micromolar

Summary

Drug discovery and development are cornerstones of the pharmaceutical industry. However, these processes are difficult, time-consuming and costly. Many new drugs fail during clinical development due to insufficient demonstration of their clinical benefit and efficacy. One of the major reasons for this failing is the lack of robust understanding of a drug's mechanism of action. In addition, phenotypic drug discovery is experiencing a revival due to disappointing results obtained through target-based drug discovery. However, phenotypic drug discovery requires challenging target deconvolution approaches to identify the cellular target of a drug. Thus, the requirement of unraveling drug-target interactions to support the drug development process is greater than ever. Many methods are available to identify drug-target interactions, although none provides a one-size-fits-all solution. Genetic approaches in particular provide strong evidence for target confirmation as they examine drug-target interactions in a living cell. Although genetic approaches for target confirmation are widely applied to simple organisms, their application to mammalian cells has been hindered for a long time. Fortunately, chemical-genetic approaches in a mammalian context have started to develop with the advent of next generation sequencing technologies, RNA interference and more recently, CRISPR/Cas-mediated genome editing. However, current chemical-genetic approaches for higher eukaryotes are often plagued by technical issues, are not well suited to probe essential genes or require costly whole-exome sequencing combined with complex bio-informatics. Therefore, the field would benefit from new approaches that can simplify drug mechanism of action studies in a mammalian context. In this thesis we explore whether CRISPR/Cas genome editing can provide such an alternative for identification of drug-target interactions in a human context.

In the first part we validated the cellular drug-target interaction of selective inhibitors of nuclear export (SINE). These small molecule inhibitors show potent anticancer activity in a variety of *in vitro* and *in vivo* models of cancer and are currently being tested in clinical trials against cancer. They have been shown to modulate the activity of the Exportin-1 (XPO1) protein through covalent interaction with XPO1's cysteine$_{528}$ residue. XPO1 is an essential protein of the karyopherin-β family of nuclear transport receptors and is required for the active nuclear export of many proteins from the nucleus to the cytoplasm. Inhibition of XPO1 function is considered a potential anticancer strategy. However, the direct causality between the observed anticancer activity of the SINE and the binding of the SINE to XPO1 has not been demonstrated. To validate this drug-target interaction we applied CRISPR/Cas9-induced homology directed repair to generate XPO1 mutant cancer cell lines that carry a serine substitution of the cysteine$_{528}$ residue. These mutant cell lines were highly resistant to treatment of the SINE on various parameters and our results validate the anticancer activity of these inhibitors is caused by selective binding to the cysteine$_{528}$ residue of XPO1.

In the second part we developed a new genetic target deconvolution method based on CRISPR/Cas-mediated genome editing. We reasoned that targeted CRISPR/Cas-induced DNA double strand breaks can be utilized for rational mutagenesis to derive gain-of-function drug resistance mutations, even in essential genes. We validated this concept using three anticancer drugs for which the drug-target interactions are well-known. We then applied this concept on a large scale to scan multiple genes simultaneously and identified nicotinamide phosphoribosyltransferase (NAMPT) as the cellular target protein of an investigational anticancer compound. We further validated this drug-target interaction by X-ray crystallography and CRISPR/Cas validation experiments. Finally, we show that the CRISPR mutagenesis approach is compatible with the class 2, type V CRISPR AsCpf1 endonuclease, increasing the resolution of the methodology. Taken together, our results highlight CRISPR/Cas-mediated genome editing provides a powerful tool for the validation and identification of drug target interactions in a human cellular context.

Samenvatting

De ontdekking en ontwikkeling van nieuwe geneesmiddelen vormen een belangrijke taak van de farmaceutische industrie. Dit proces is zwaar, tijdrovend en kostelijk en veel nieuwe medicijnen geraken niet door de klinische ontwikkelingsfasen omdat ze onvoldoende activiteit vertonen. Een belangrijke oorzaak voor dit falen wordt gezien als onvoldoende bevestiging van het doelwit en werkingsmechanisme van het kandidaat medicijn. Hierdoor is het ontrafelen van het doelwit en werkingsmechanisme van een medicijn een belangrijke stap in het ontwikkelingstraject van een nieuw medicijn. Er zijn verschillende methoden voor de ontrafeling van het doelwit van een medicijn, maar geen enkele biedt een algemeen toepasbare oplossing. Niettegenstaande bieden genetische methoden een sterk platform voor doelwitidentificatie omdat zij toelaten interacties tussen medicijnen en doelwitten in een cellulaire context te bestuderen. Met de ontwikkeling van "next generation sequencing", RNA-interferentie en recentelijk CRISPR/Cas-gemedieerde genoommodificatie, is het immers mogelijk geworden om genetische methoden op grote schaal toe te passen in menselijke cellen. Sommige van deze methoden kampen echter met technische problemen, of kunnen niet gebruikt worden om essentiële genen te onderzoeken, terwijl weer andere dure sequencing gekoppeld aan complexe bioinformatica vereisen. Hierdoor heeft het veld nood aan alternatieven die toelaten het doelwit en werkingsmechanisme van een nieuw medicijn makkelijk en goedkoop te identificeren. In deze thesis onderzoeken wij of CRISPR/Cas-gemedieerde genoommodificatie een meerwaarde kan bieden voor de validatie en identificatie van interacties tussen doelwitten en kandidaat medicijnen in een menselijke context.

In het eerste deel valideren wij de interactie van klein molecule kandidaat geneesmiddelen met exportine-1 (XPO1). XPO1 is een essentieel eiwit en faciliteert het actieve transport van eiwitten vanuit de celkern naar het cytoplasma. Deze eiwitten omvatten onder andere transcriptiefactoren, tumorsuppressoren en groei-regulatorische eiwitten. Een correct nucleair-cytoplasmatisch transport is dus essentieel om de cellulaire homeostase te bewaren. De XPO1 inhibitoren vertonen een sterke anti-kankeractiviteit in verschillende laboratorium- en dierproeven en worden momenteel onderzocht in klinische studies tegen kanker. Verder is aangetoond dat deze geneesmiddelen de functie van het cellulaire XPO1 eiwit verhinderen door covalent te binden op het cysteine$_{528}$ residu van XPO1. Nu was het niet bewezen dat de anti-kankeractiviteit van de XPO1 inhibitoren veroorzaakt wordt door verstoring van de XPO1 functie. Om dit te bewijzen pasten wij CRISPR/Cas-gemedieerde genoommodificatie toe om in kanker cellijnen het cysteine$_{528}$ residu in XPO1 te vervangen door een serine. Deze mutante cellijnen waren zeer ongevoelig aan de XPO1 inhibitoren, wat bewijst dat de anti-kankeractiviteit van deze middelen afhankelijk is van het cysteine$_{528}$ residu in XPO1.

In het tweede deel ontwikkelen wij een nieuwe genetische methode om het cellulaire doelwit van een kandidaat geneesmiddel te identificeren via CRISPR/Cas-gemedieerde mutagenese. Wij redeneerden dat CRISPR/Cas-gemedieerde DNA-dubbelstrengbreuken het mogelijk maken om doelgerichte "gain-of-function" resistentie mutaties te bekomen in essentiële genen. Om dit concept te valideren pasten wij CRISPR/Cas-gemedieerde genoommodificatie toe op drie anti-kankermedicijnen waarvan het cellulaire doelwit gekend is. Na deze validatie pasten wij dit concept op grote schaal toe waarbij meerdere genen tegelijk afgetast werden op resistentiemutaties om nicotinamide-fosforibosyltransferase (NAMPT) als het cellulaire doelwit van een kandidaat geneesmiddel te identificeren. Deze bevinding werd verder gevalideerd via X-stralen kristallografie en CRISPR/Cas-gemedieerde genoommodificatie. Daarnaast toonden wij ook aan dat de methode compatibel is met het klasse 2 type 5 AsCpf1 CRISPR endonuclease. Samen laten onze resultaten zien dat CRISPR/Cas-gemedieerde genoommodificatie een krachtig hulpmiddel biedt voor de validatie en identificatie van interacties tussen medicijnen en hun doelwit.

CHAPTER I

General introduction

1.1 The drug discovery and development process

Drug discovery and development are major cornerstones of the pharmaceutical industry. The entire process is complex, difficult and infamously expensive (±1-3 billion US$ for a single drug) and requires multi-disciplinary research teams to work in close collaboration for long periods of time[1-6]. It first aims to discover candidate drugs that can modulate a disease or phenotype of interest. After this initial drug discovery phase, it aims to further develop the candidate drugs into safe and efficient therapies for which side-effects can be predicted and controlled.

As with any project, the first step in drug discovery and development involves a project-planning phase. In this phase, the researchers evaluate the commercial potential, medical need and relevance of the problem they wish to solve and what resources and infrastructure are required to succeed. Once a project plan has been approved, the drug discovery phase can start. Two drug discovery approaches are generally distinguished, namely phenotypic and target-based drug discovery. Both involve screening a large collection (10^3-10^6) of molecules in a set of relatively simple and high-throughput assays to nominate a few hit molecules that modulate a disease or phenotype of interest[3, 7, 8]. Historically, nature itself provided humankind with a wide variety of molecules[9-11]. However, in addition to these natural compounds, many new molecules have now been and still are being synthesized with increasing support from virtual screening and computer-aided drug design[3, 12-14]. Currently, the drug discovery phase still takes a few years and the success rate is relatively unpredictable and target dependent[2-4].

After completion of the drug discovery process, nominated and optimized lead hit molecules are studied extensively and characterized for toxicity, metabolism and safety in a more biologically relevant context, typically involving cutting edge biological model systems and live animals. After the candidate molecules have successfully gone through this preclinical development stage, they are ready to be tested for the first time in humans during the clinical studies. These studies are expensive, time consuming and divided into multiple phases in which the number of humans treated with the candidate drug steadily increases. Once a drug has passed through all clinical phases and shows efficacy and efficiency for treating a disease, it is ready to be registered and submitted for approval to the appropriate agencies. Once approved, the drug can be released to the market and sold to patients and clinicians. However, during this marketing phase, it is often required or recommended to monitor how patients respond to the candidate drug to keep track of previously undetected adverse events and issues related to the drug. Of note, the complete drug development process does not always have to be as linear as shown here (**Fig. 1**) and especially the drug discovery and preclinical development phases often intertwine[15-17]. In addition, the same candidate drug can be tested against multiple subtypes of a disease or even against different diseases all together and it is possible that multiple preclinical and clinical studies are being performed at the same time for a given candidate drug.

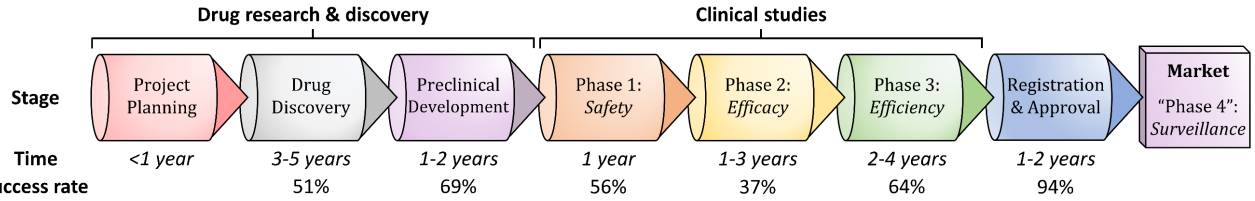


Figure 1: *simplified overview of the different stages in the drug discovery and development process.* After planning the project, a large collection of molecules is tested during drug discovery to identify drug candidates. These candidate molecules are then characterized for safety and efficacy in biological and animal models (preclinical development) before being tested in humans (clinical studies). During the clinical studies, the drug is first tested for safety and toxicity in a small group of humans and the minimal toxic dose is determined using dose-escalation studies. During phase 2 and 3, larger groups of patients are tested for efficacy and efficiency against the disease of interest. If successful, the drug can be submitted for approval by the medicine agencies, and once approved, the drug can be released to the market. Each success rate indicates the chance for success at each stage independently. Combining these rates gives an overall success rate of 4.3% for the entire process. Success rates were obtained from[4, 18].

1.1.1 Phenotypic drug discovery

A typical approach used in the drug discovery process is known as phenotypic drug discovery (**Fig. 2**). This approach aims to identify molecules that can modulate the progress of the disease of interest without focusing on the underlying molecular mechanism of action at first[3, 19-21]. It relies on the development and selection of easy to use, predictive and scalable biological assays that generally consist of a cellular system representative of the disease of interest[20, 22, 23]. A large set of random molecules are then screened in the selected assays for the desired phenotypic change. After screening, hit molecules showing desired bioactive properties are selected using predefined criteria. The goal then shifts to determine and validate the cellular and molecular mechanisms underlying the biological activity of the hit molecules to determine the drug's selectivity and target, a process called target deconvolution. The next step, termed hit-to-lead generation, involves optimization of the hit molecules using medicinal chemistry approaches such as pharmacophore modelling and structure activity relationship studies to increase the activity, stability and drug likeness properties of the hit molecules. These optimized hit molecules are then screened and tested in secondary follow-up and counter-screen assays and revalidated for their underlying mechanism of action (target validation).

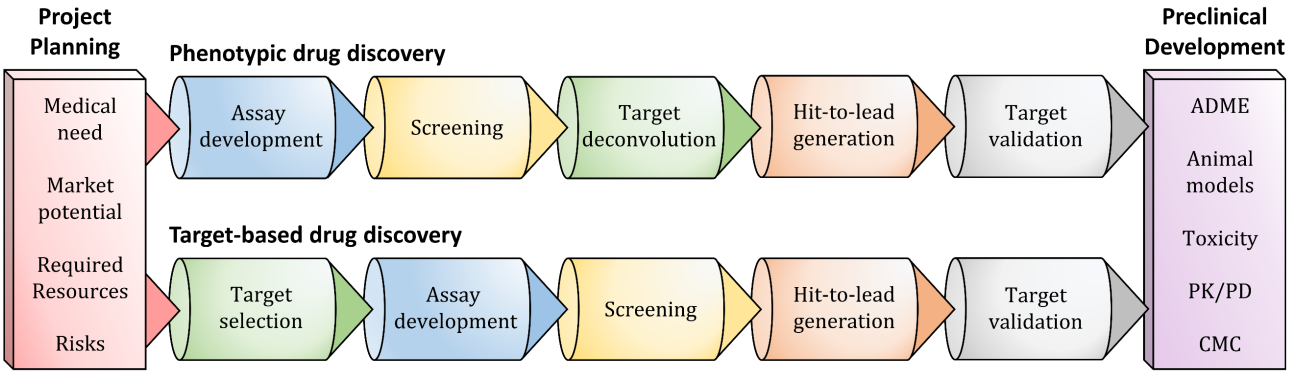


Figure 2: *overview of the different steps in early drug discovery and development.* Following the planning phase, two different drug discovery approaches are generally distinguished. Phenotypic drug discovery aims to identify hit molecules by screening many compounds in carefully selected complex biological assays representative of the disease. Target deconvolution is then performed to identify the molecular mechanism of action underlying the bioactivity of the hit molecules. Target-based drug discovery starts with the selection of a specific target or molecular mechanism of interest and then aims to identify hit molecules that modulate this process through screening. After identification and validation of the lead molecules, both drug discovery approaches end up in the preclinical development phase. **ADME**: absorption, distribution, metabolism, excretion. **PK/PD**: pharmacokinetics /pharmacodynamics. **CMC:** chemistry, manufacturing and control.

1.1.2 Target-based drug discovery

A second approach, known as target-based drug discovery, follows a reverse process as phenotypic drug discovery (**Fig. 2**)[2, 3, 24]. It first aims to identify, dissect and validate a single molecular process involved in a disease (target selection). After selection of a molecular target, a large set of molecules are selected or newly designed based on the chemical properties of the chosen molecular target and then screened for modulating activity against the molecular process in a confined and controlled biochemical assay. Hit molecules are then nominated based on predefined criteria and tested in more complex biological models relevant to the disease of interest. The researchers then perform hit-to-lead generation on the most promising bioactive hit molecules using medicinal chemistry and perform follow-up validation and cross-screening using secondary assays. Finally, the researchers validate whether the bioactivity of the optimized lead molecules is still caused by selective modulation of the molecular mechanisms defined at the start of the process (target validation).

1.1.3 Preclinical development

Following the drug discovery process, the most promising drug candidates are selected for further chemical optimization and testing during preclinical development[2, 3, 7, 14]. Experiments at this stage focus on determining whether treatment of humans with these molecules is expected to be feasible, effective and safe. However, as it is not ethically warranted to test drug candidates immediately in humans or even animals, early characterization of their chemical characteristics and metabolism is performed in *in vitro* biochemical and biological assays by determining absorption, distribution, metabolism and excretion rates (ADME)[25, 26]. The most promising candidates are then administrated to living organisms such as fruit flies, mice, dogs, pigs or even primates[2, 3]. The goal of these *in vivo* models is to obtain a clear profile of how a living organism is affected by the drug (pharmacodynamics, PD) and how the drug is altered by the organism (pharmacokinetics, PK). For this purpose, the effect of drug treatment on the organism is characterized by determining target binding, toxicity, carcinogenicity, tolerability and efficacy profiles and the occurrence of adverse events and side effects. In addition, the rate of distribution, metabolism and excretion of the drug by the model organism is determined.

Once a clear profile has been drafted, the researchers must decide whether to advance their main candidate drugs to the next stage, human clinical trials. However, before this stage can start, the researchers are typically challenged with another task. They must determine and find a way to scale up the synthesis process of their candidate drug from producing a few milligrams, to manufacturing grams, kilograms or even tons of drug. In addition, the researchers need to determine how they plan to administer, distribute and package the candidate drug. For example, as capsules, tablets, or injectable formulations. Together, these steps are known as chemistry, manufacturing and control (CMC). Finally, before phase 1 clinical studies can start, the researchers are required to submit an investigational new drug (IND) application to the regulatory authorities. This IND contains the information on preclinical safety and toxicity, *in vivo* pharmacodynamics and pharmacokinetics profiles, CMC information and typically also requires a first draft containing protocols for the proposed clinical studies.

1.1.4 Human clinical trials

Human clinical trials are lengthy, tightly controlled, expensive and a crucial phase of the drug development process[4, 5, 27, 28]. They are generally subdivided into four phases and aim to determine whether treatment of human beings with the candidate drugs is safe, feasible and effective (**Fig. 3**). All clinical trials are carried out in close collaboration with clinical centers and healthcare professionals and require approval from the medical agencies. The first phase, also known as “first-in-man” studies, is often the first time live human beings are treated with the candidate drug. Phase 1 trials are designed to include a small amount of approximately 10-100 (healthy) volunteers at a single or a small number of clinical centers and aim to identify whether treatment of humans with the candidate drug is safe. Another important goal of phase 1 trials is to determine toxic effects and the maximum dose that can be tolerated before adverse effects become problematic. Phase 1 trials typically last for 6-12 months and 50-70% of the candidate drugs entering a phase 1 trial pass successfully[18, 27, 29].

Once phase 1 has been passed and a dosing regimen has been established, phase 2 clinical trials can commence. Phase 2 trials are designed to assess whether the candidate drug has the desired bioactivity or effect and can modulate a disease of interest in humans. These trials are a critical point in the drug development process with 50-80% of candidate drugs failing during this phase[18, 29-32]. They are carried out in a limited number of clinical centers, typically last for 1-2 years and include 10-500 patients diagnosed with the disease of interest[27, 33]. They focus on specific manifestations or subtypes of a disease and multiple phase 2 trials with different groups of patients can be carried out simultaneously. Besides assessing bioactivity and efficacy, phase 2 trials also provide additional information on the occurrence of side effects and adverse events and help researchers to further determine the optimal dose.

Phase 3 studies are carried out after extensive deliberation of drug safety and efficacy profiles obtained in phase 1 and 2[4, 27]. The major goal of a phase 3 trial is to provide a definitive assessment of whether treatment with the candidate drug is effective in a large clinical setting and provides an overall benefit to patients by comparing the candidate drug to the gold standard treatment currently available. In addition, phase 3 trials aim to identify and estimate the incidence of common adverse events. Due to these goals, they require strong statistical power and are carried out with large patient groups of 500-5000 patients spread across multiple clinical centers[27, 34]. Patient groups are selected based on the diagnosis of a specific variant of a disease and are generally more diverse than in phase 1 and 2 studies. Phase 3 studies can last for multiple years, especially in the case of chronic diseases, and are known to be expensive. Moreover, they challenge the logistical and organizational administration of the company developing the candidate drug, as they now must deliver the drug in large quantities to patients at multiple clinical centers spread across the world. However, due to extensive deliberation and assessment of the available data surrounding the candidate drug prior to starting a phase 3 study, they tend to have a higher success rate than phase 2 studies.

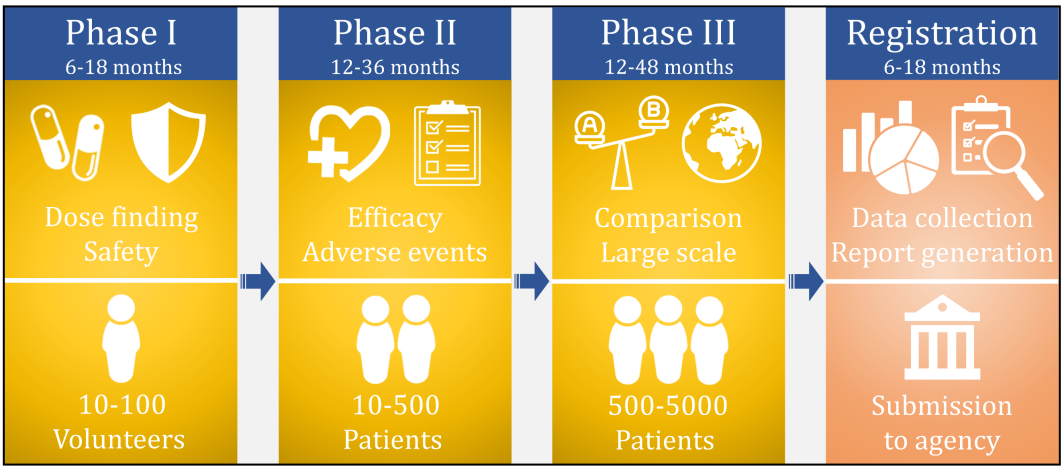


Figure 3: *overview of the different phases for human clinical trials.* During the first phase, a small group of (healthy) volunteers are treated with the candidate drug to determine drug safety and dosing. During phase 2 trials, the drug is administrated to real patients and examined on its potential to cure disease. In phase 3 studies the patient groups are extended, and the candidate drug is compared to the standard of care already available to patients.

1.1.5 Registration and post-approval

Once a candidate drug has completed phase 3 trials, the drug can be registered at the medicine agencies, such as the North-American Food and Drug Administration (FDA) and the European Medicines Agency (EMA)[27]. These will evaluate and review the profile and results of the candidate drug to determine whether the benefits outweigh the drug's known and potential risks for the intended population. These agencies may ask the drug's developer to carry out additional studies to better characterize the safety profile and related adverse events of a candidate drug before giving their approval. Once approved, the drug can be marketed and sold to patients and clinicians. Typically, pharmaceutical companies continue surveilling the use of their drug after approval and may carry out post-marketing studies, or "Phase 4" studies. The aim of these studies is to identify the occurrence of less common adverse events and to evaluate the cost/effectiveness ratio of the drug. In addition, patients and clinicians are able and encouraged to voluntarily report any problems or adverse reactions related to the newly approved drug to the agencies.

1.1.6 The importance of target validation for drug development

The therapeutic effect of most clinical drugs is achieved through modulation of the function of one or a few target proteins in a cell or tissue. This modulation is typically achieved through direct binding of the drug at a functional site of the target protein and can either lead to activation or inhibition of the protein's function. The extent of this engagement between the target protein and the drug dictates the drug's efficacy, and adverse events often occur due to excessive binding of the drug to the target or by engagement of the drug with other proteins[35]. As such, having solid knowledge on the molecular target(s) of a candidate drug helps to predict possible side effects and the validity and tractability of the envisioned therapy[3, 36-39]. A detailed understanding on how a drug interacts with its target protein also supports hit-to-lead optimization and biomarker development. If strong biomarkers related to the drug's target can be uncovered, these can help to predict which patients will or will not respond to the candidate drug, facilitating the correct selection of patient groups for clinical studies.

Although the total number of drugs approved each year is increasing, the approval of candidate drugs that target novel molecular mechanisms for which no existing drugs have been approved before has stalled[4, 40]. Failure to reach approval is traditionally attributed to four major issues related to the candidate drug: (1) poor pharmacokinetics, (2) high toxicity, (3) insufficient efficacy and (4) case-dependent reasons such as administrative or regulatory issues[4, 6, 30-32]. In the last 20 years, the industry has made significant progress with reducing the attrition rate due to poor pharmacokinetics by improving existing *in vitro* ADME assays, developing *in silico* prediction models for *in vivo* ADME and by understanding the translational power of *in vivo* pharmacokinetic studies carried out with model organisms better[25, 30]. The attrition rate caused by dose limiting toxicity issues is also expected to lower significantly due to the incorporation of toxicity studies in the early drug discovery and preclinical development phases, although this effect remains to be investigated further[41, 42].

Many drugs are suggested to fail, especially during phase 2 clinical studies, due to a lack of efficacy[4, 18, 29, 30, 32]. A major reason for this failing is insufficient target validation as many target-disease associations derived from preclinical efficacy data obtained with *in vitro*, *in cellulo* and *in vivo* model systems do not translate well to human patients. Interestingly, in addition to the lack of translational power of preclinical models, differences between the reasons for failure between candidate drugs identified through phenotypic and target-based drug discovery approaches can be distinguished. Candidate drugs discovered through target-based discovery often fail in clinical stages because the initial selection of the target was based on limited data. This can lead to the selection of a target that might not be as relevant to the disease of interest as initially was believed. Another reason for the failure of these drugs is found in the insufficient profiling of the (poly) pharmacology of the early hit molecules as profiling studies tend to bias around the initially selected target protein. This can lead to the association of wrong target-phenotype relations and to hit-to-lead generation based on flawed pharmacology principles. Although target-based discovery approaches have dominated the drug discovery field since the late 80s, a marked resurgence of phenotypic screening has been observed over the last decades due to practical improvements and more importantly due to the disappointing approval rate for drugs obtained by target-based discovery[21, 24, 43]. However, phenotypic approaches require target deconvolution, which is a very challenging process. Indeed, for some drug candidates the target is never identified and therefore many candidates discovered by phenotypic screening end on the shelf. Moreover, even if a candidate's target has been identified, it needs to be validated whether this drug-target interaction relates to its bioactivity to prevent assuming a candidate drug shows bioactivity by modulating a target of interest, although in reality they do not. Thus, to minimize the failure rate of new drugs in development, careful target identification and validation should be carried out in the early stages of the drug discovery and development process.

1.2 Methods for target deconvolution

Given the importance of understanding the mechanism of action of a candidate drug and its corresponding molecular target for the success of any drug discovery and development program, many target deconvolution (or target identification) approaches have been developed[37-39, 44-47]. These methods can be broadly categorized into three groups (**Fig. 4**). Direct biochemical methods focus on identifying biochemical interactions between a candidate drug and cellular proteins or molecules. Genetic interaction methods focus on identifying genes relevant to the efficacy of the candidate drug and methods based on computational interference aim to integrate patterns and information present in reference datasets or databases containing information on previously studied drugs to predict interactions for a given candidate drug. However, every method has its flaws and strengths and hardly any successful target deconvolution project relies on a single method. Instead, solid and successful target deconvolution depends on the clever use and integration of multiple and complementary approaches. Most of these methods aim to unravel interactions between cellular proteins and drugs, but, although most bioactive drugs act by binding to select cellular proteins, it must be noted some drugs, for example DNA crosslinking agents and photochemotherapeutics[48, 49], do not necessarily exert their bioactivity by binding cellular proteins. Moreover, due to the complex biology of eukaryotic organisms, the identification of the cellular target of a bioactive drug candidate remains highly challenging in these organisms. What follows here is an overview of the different methods available for target deconvolution.

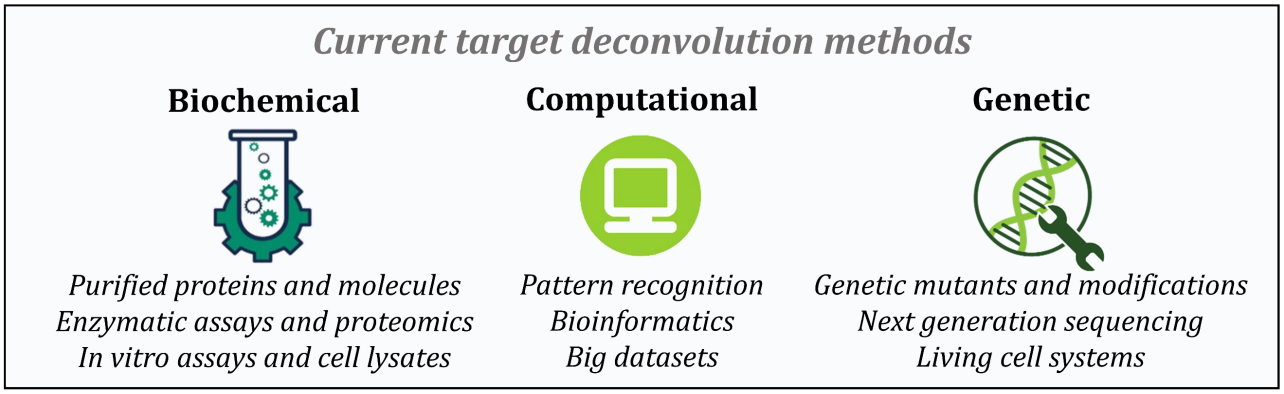


Figure 4: *overview of the three categories of methods for target deconvolution.*

1.2.1 Direct biochemical approaches to target deconvolution

Direct biochemical approaches have laid the pioneering groundwork for target deconvolution. These methods typically rely on *in vitro* systems to study the direct binding between target proteins and candidate drugs[39, 45, 46, 50-54]. Affinity-based biochemical methods have emerged as the biochemical standard for unbiased target deconvolution. In brief, a chemically optimized derivative of the candidate drug is immobilized on a cross-linked polymer or gel-like matrix. Using this matrix containing the immobilized drug, it is possible to extract proteins with a strong binding affinity to the drug from a cell lysate containing all intact cellular proteins and protein complexes. After washing away the non-binders and proteins with low affinity interactions from the matrix, the remaining proteins bound to the immobilized reagent can be eluted by altering parameters such as pH, charge or salt concentration, or by competition with excessive candidate drug. The eluted proteins can then be identified using proteomic approaches. The identification of peptides and proteins in the eluted fraction used to present a real challenge, but with the increasing availability of genetic information and the continuing development of (quantitative) protein mass spectrometry, identification of the eluted proteins has become a straightforward process[50, 54-57].

A major challenge for affinity-based methods remains to determine a way to immobilize the candidate drug to the matrix without altering the drug's reactivity and chemical properties, as alteration of a drug's properties may lead to the capture of unrelated target proteins and could abolish the interaction of the drug with its original target proteins. In addition, as washing steps can be quite harsh, the capture of weak interacting or low-abundant proteins can be challenging. Moreover, as many candidate drugs are unique, multiple chemical configurations need to be tested and it remains difficult to predict upfront which strategy will lead to a successful experiment[50, 54, 58]. The high risk of identifying false-positive target proteins due to background interactions presents another challenge for affinity-based approaches. To circumvent this problem, a control experiment is generally required to characterize the background signal. For example, one can make use of inactive derivatives of the candidate drug, incubate a matrix or gel that does not contain the immobilized drug or preincubate the cell lysate with a soluble and mobile form of the candidate drug to determine background signals.

Besides immobilizing the compound to a matrix, it is also possible to link the drug directly with an affinity handle such as biotin, a fluorophore or a click chemistry handle that can be used to capture the drug in later stages[45, 46, 51, 59-62]. These approaches tend to work with covalent inhibitors and allow treatment of living cells if the modified drug can cross the cell membrane. As this activity-based approach relies on covalent interactions it is possible to perform stringent washing steps during sample preparation to limit the amount of background signal, although even in these cases background signals can still be problematic.

As it is widely recognized chemical optimization of candidate drugs for affinity-based approaches is troublesome, new approaches that circumvent the requirement of immobilizing compounds are actively being developed. One of these methods, termed Drug Affinity Responsive Target Stability (DARTS), utilizes quantitative mass spectrometry to determine the change of a protein's susceptibility to proteolytic degradation when bound by a candidate drug[63-65]. Another label-free method termed target identification by chromatographic co-elution (TICC) has recently been reported and utilizes *in vitro* based liquid chromatographic separations of cell lysates in the presence of drug to monitor the interaction of cellular proteins with a candidate drug[66]. In addition, thermal proteome profiling (TPP) of denaturation curves of proteins present in a cellular protein lysate in the presence or absence of a candidate drug using quantitative protein mass spectrometry has been shown to be able to support target deconvolution[67-72]. However, the alteration of the physicochemical properties of a protein by binding to a candidate drug differs strongly between individual proteins and thus not every drug binding event will lead to detectable stabilization of a full-length protein. In addition to these new approaches, other methods are expected to be developed and it will be interesting to see how this field will continue to evolve.

1.2.1.1 Biased biochemical approaches

When strong evidence is already available to pinpoint potential target proteins, biased biochemical approaches can be used for target validation and deconvolution. For example, when the target protein is likely to be an enzyme, the activity of a candidate drug can be profiled against a panel of predetermined enzymes, which are nowadays commercially available in an optimized ready-to-use format or service[50, 59, 73-76]. If a candidate drug can be labelled fluorescently, it can be added to a panel of purified proteins in a microarray format to identify which proteins captured the fluorescent molecule. Like the proteome-wide thermal shift profiling methods, thermal/chemical shift assays using individually purified proteins can also be used to determine drug-binding properties of candidate drugs to a select group of proteins[77-79]. Next to these techniques, surface plasmon resonance can be used to characterize differences in the reflection of polarized light shined on a metal surface embedded with purified proteins in the presence or absence of a candidate drug[80]. Engagement of the drug with an embedded protein alters the profile of the reflected light, which enables the identification of interactions.

Many other methods to determine target engagement of a candidate drug with a select set of proteins can be distinguished and include homogeneous time resolved fluorescence (HTRF), hydrogen-deuterium exchange (HDx) and circular dichroism-based (CD) methods as well as size-exclusion chromatography for target identification (SEC-TID) and *in vitro* ribosome-display methods such as parallel analysis of translated ORFs (PLATO)[81-86]. As biased biochemical methods tend to be limited to a small amount of proteins, they are used mainly for target validation and measuring target engagement. Thus, these methods are ideal for setting up a target-based drug discovery program or guiding medicinal chemistry approaches to optimize the activity and selectivity of a candidate drug. Many biased and unbiased biochemical assays tend to rely on the use of purified proteins and therefore lack the complexity and biological context of a cell or tissue. In addition, purified proteins are often used at artificial concentrations not representing physiological relevant expression levels, further compromising the biological relevance of the findings. Finally, as not all proteins can be purified easily or require complex scaffolding structures such as an intact cell membrane, the potential and relevance of some biochemical approaches, including unbiased affinity-based approaches, can be constrained.

1.2.2 Computational approaches to target deconvolution

Computational approaches rely on pattern recognition to integrate results obtained from different experiments or contained in reference databases on existing drugs[39, 45, 87]. They tend to unbiased as they often rely on publicly available data and can be used as a standalone approach for target deconvolution but will require strong experimental validation. As such, these methods shine when used in combination with direct biochemical or genetic methods.

A well-known class of computational methods can be broadly defined as the profiling approaches. These methods rely on comparing biological profiles in the form of functional readouts of drug action on targets, cells or even organisms[88-90]. They depend on the assumption that drugs targeting the same target protein or having a similar mechanism of action will show a similar behavior across the biological assays used to obtain the readout. A seminal example for profiling methods was set by the US National Cancer Institute “NCI-60” cancer cell line dataset[91]. With this dataset, researchers were able to link protein targets using their expression levels to the sensitivity of different anticancer drugs[92]. Other early profiling efforts succeeded on comparing genome-wide expression profiles in yeast, rat and human cells treated with different reference compounds to the expression profile of cells treated with an investigational drug to infer possible drug-target interactions[93-97]. The effect of a candidate drug on a battery of biological assays or a list of cellular parameters can also be compared to the effect of a set of reference compounds to infer the target of a candidate drug[89, 98]. Image-based studies focusing on microscopic readouts obtained with high-throughput microscopy have been performed to cluster different classes of compounds and to nominate potential drug-target interactions[99-101]. Profiling of drug modulatory properties on cell death induction by another drug has been used to identify the microtubule network as the target of three uncharacterized anticancer drugs[102].

In addition to these profiling approaches, one can also compare the biochemical structure and properties of a drug to the structure and properties of other drugs and ligands with a known mechanism of action. Various methods for ligand-based prediction approaches exist[103-106]. The data mining similarity ensemble approach (SEA) uses known protein ligands to group target proteins into different classes. Proteins with similar ligands and chemical mode of ligand-binding group together and using these protein groups, SEA was able to predict off-target and on-target proteins for some drugs[107]. Prediction of activity spectra for substances (PASS) aims to predict biological activities for a specific compound by the structural chemical formula of the drug alone. It incorporates a detailed analysis of the structure-activity relationships of many known small molecules and uses this data to predict the activities of the investigated drug[108, 109]. Machine learning techniques have also been developed to infer the

target of a drug[110-113]. A specific example of such a method utilized protein domains instead of full length proteins to associate substructures of drugs to targets and supported target deconvolution using an affinity-based biochemical method[114]. However, all these methods are limited by the availability of reference datasets or collections and are not always unbiased, limiting their potential. Finally, structure-based methods for target deconvolution attempt to use protein crystal structures to dock and fit drugs unto target proteins[115-118]. As these methods require the availability of a protein crystal structure, they lack the power to perform unbiased and holistic target deconvolution. However, these methods, when successful, provide important biochemical information and allow for powerful hit-to-lead generation.

1.2.3 Genetic approaches to target deconvolution

Functional genetic approaches rely on manipulating the DNA or RNA in a cellular context and linking these genetic perturbations to biological phenotypes. Typically, two genetic screening approaches are distinguished. Forward genetic screens rely on isolating individuals showing a desired phenotype from a genetically diverse population. Such diverse populations can exist naturally but can also be obtained by large-scale manipulation of genetic elements. When the desired individuals have been isolated, the goal switches to identify which genetic mutation or alteration causes the desired phenotype. Reverse genetic screens on the other hand aim to alter a specific genetic element upfront and then identify the phenotypes caused by this specific genetic alteration. Due to the ease of manipulating DNA and RNA inside living cells, genetic approaches provide a relatively easy to use and powerful tool for target deconvolution[37-39, 45, 119]. Genetic target deconvolution approaches generally represent forward genetic screens as target deconvolution is performed because the target gene or protein of a drug is unknown. However, although reverse genetic approaches are not very useful for unbiased target deconvolution, they do provide a powerful framework for target validation[120, 121].

1.2.3.1 Yeast: a prototype model system for chemical-genetic interaction screening

A central concept in establishing chemical-genetic interactions follows that the expression and concentration level of a target protein inside a cell correlate directly to the sensitivity of this cell to a drug acting on this target protein[35, 122]. The establishment of this principle is strongly supported by early yeast genetics and has been used to identify the target protein of candidate drugs ever since[123]. For example, yeast strains carrying loss-of-function mutations in specific genes were more sensitive to agents targeting proteins related to these genes[124, 125]. Likewise, overexpression of specific genes using multicopy plasmids can desensitize yeast strains to certain drugs[123, 126]. Based on this principle, different libraries of isogenic heterozygous yeast mutants containing loss-of-function mutations in one allele and a wild-type sequence in the other have been generated[127-132]. Some of these libraries span almost every gene and when each unique mutant is barcoded, it becomes possible to perform large-scale arrayed or pooled genetic screens[133, 134]. Heterozygous loss-of-function mutations that sensitize or desensitize these yeast strains to drug treatment can then by identified by sequencing and quantification of the barcodes present before and after treatment with a drug of interest. This process, termed haploinsufficieny profiling (HIP), facilitates the identification of chemical-genetic interactions and has been used to identify the target of different classes of drugs, including antifungal, anticancer and metabolic agents[127-132]. However, not all drug-target interactions show haploinsufficieny profiles and these studies, although they can provide additional information on the mechanism of action of a candidate drug, are not always able to identify the direct cellular target of a drug[128, 132].

Similar methods to decrease the expression or functional activity of genes in yeast strains have been reported and rely on inducible gene promotors, tagging of the untranslated 3' end of genes with destabilizing markers (decreased abundance by mRNA perturbation, DAmP) or on isolating temperature-sensitive alleles[135-140]. These methods allow lowering of

gene expression in a more continuous fashion and can inhibit expression more than the typical 50% when using heterozygous loss-of-function mutations. Profiling these altered strains for sensitivity to a candidate drug also allows for the identification of chemical-genetic interactions[136, 137, 141].

These knockdown approaches often lead to difficult to interpret results which typically include only a handful of weak potential target genes. For this purpose, researchers have turned to complete loss-of-function yeast mutants, which are created by homozygous deletions in both alleles of a gene[37, 38, 142-144]. However, a major drawback of these homozygous deletion profiling (HOP) methods is that they are unable to probe essential genes, as complete loss-of-function of an essential protein will kill the cell. In addition, results obtained with homozygous deletion profiling do not always reveal the direct cellular target protein of a candidate drug. Instead, they often nominate accessory proteins that are involved in drug sensitivity and down-stream mechanism of action of the candidate drug.

Besides a lowered or abolished expression, an increased expression of a gene can also alter drug sensitivity of a cell. Libraries of yeast strains carrying additional copies of a gene have been used to screen for genes modulating drug responses. This multicopy suppression profiling (MSP) approach has led to the identification of the target of different drugs and has the potential to identify the direct target of a drug[145-147]. However, as both overexpression and loss-of-function profiling are prone to identify false positives and negatives, a combination of overexpression profiling together with HIP or HOP profiling has been suggested to lead to better results. Indeed, a proof-of-concept study has shown such a combination in yeast leads to more sensitive and powerful target deconvolution[148].

As the bioactivity of many drugs is exerted through physical interaction of the drug with its target protein, alteration of the drug-binding sites on the target protein should block this interaction and abolish the bioactivity of the drug. This concept forms the basis for a genetic target deconvolution approach that relies on the isolation of point mutations that abolish drug binding and confer drug resistance in a cellular context. These mutations typically do not alter the function and expression level of the protein. This concept has been applied in a seminal study using a yeast model system to screen for the cellular target of rapamycin[149]. Another well-known example is the identification of the ARF-GDP-Sec7 domain protein complex as the target of brefeldin A[150]. Importantly, although every cellular system is subject to a background rate of mutation, mutations are often induced using chemical mutagens such as *N*-ethyl-*N*-nitrosourea (ENU) or ethylmethane sulfate (EMS), through UV radiation, by expressing a library of artificial variants or by lowering the inherent DNA repair capacity of the organism[151, 152].

1.2.3.1.1 Limitations of yeast model systems

Importantly, although yeast model systems are easy to work with, were used to lay down the groundwork for chemical-genetic studies and allow for high-throughput screening[37, 38], they have limitations in their applicability for target deconvolution of drugs intended for humans. Unlike humans, yeast cells have a cell wall which can severely impact the uptake of a drug. In addition, yeast cells express many protein pumps, of which some are yeast-specific, that allow for rapid efflux of drugs entering the cell. This can lead to strong intrinsic drug resistance which is not related to a drug's target protein or cellular mechanism of action. Moreover, yeast cells do not form complex multicellular tissues and lack detailed differentiation states, limiting their representative power of the biological complexity of a human cell. Finally, their genome size is smaller and less complex than that of higher eukaryotes and as a result, many human genes are not well represented by in yeast model systems.

1.2.3.2 Chemical-genetic screens in human cells

With the completion of the human genome project and the advent of next-generation sequencing, the foundation for genetic screening in human cells had been laid down in the early 2000s[153-155]. However, new technologies were needed to effectively perform genetic screening in human cells as most of the techniques used in yeast were not directly transferable to human cell culture. For example, genetic methods relying on backcrossing of different strains are not compatible with human cell culture. Luckily, the past 20 years have seen a rapid development of functional genomics tools for human cells.

1.2.3.2.1 RNA interference

The discovery of RNA interference (RNAi) in eukaryotic cells paved the way for prototype genetic loss-of-function screening in a human cellular context[156-161]. Briefly, RNAi relies on short noncoding RNA fragments (siRNA/shRNA) that are complementary to a unique RNA sequence present in the mRNA of an expressed gene to specifically degrade the targeted mRNA by hijacking the endogenous RNA-induced silencing complex (RISC) (**Fig. 5**)[159]. This targeted degradation of mRNA then leads to a reduced expression of the target gene. Large scale screening with RNAi typically involves the use of lentiviral particles which can be used to transduce almost any mammalian cell type *in vitro*. Moreover, each shRNA can be designed to be unique, allowing each shRNA sequence to be used as barcode.

However, although RNAi provides a powerful methodology for reducing protein expression levels, the mRNA for different genes is expressed and processed at different rates. This makes it difficult to modulate the expression of every single gene with similar power during large-scale RNAi experiments. Moreover, when RNAi screening is performed in complex or different biological systems, this problem is exaggerated as the expression level of the same gene can vary greatly between cell types. It has also become apparent that RNAi reagents tend to show high off-target interference, resulting in degradation of mRNA that was not meant to be targeted, which can contribute to the discovery of false positive hits[158, 161, 162]. These shortcomings have in part been addressed by a better understanding of the machinery underlying RNAi, by optimizing RNAi reagents and delivery and by using multiple RNAi reagents targeting a single mRNA molecule[163-165]. In addition, better analysis software has been developed to tackle the issues with interpreting RNAi data[166, 167].

Due to these improvements, it is possible to perform genome-wide RNAi screens in human cells for target deconvolution. For example, high complexity genome-wide RNAi screening was used to identify nicotinamide phosphoribosyl transferase (NAMPT) as the direct target protein of the anticancer compound STF-118804[168]. Another RNAi-based screen revealed that ISRIB, a compound found to enhance cognitive memory, acts by inducing dimerization of eIF2B4[169]. However, similar to HIP and HOP profiling in yeast, RNAi screening can lead to the identification of genes that are indirectly involved in a drug's mechanism-of-action. For example, genome-wide RNAi screens were used to map PTPN1, NF1, SMARCB1, and SMARCE1 as important proteins for sensitivity to the anticancer compound imatinib and identified that cellular sensitivity to Nutlin-3, an MDM2 inhibitor, is dependent on TP53BP1[170, 171]. Furthermore, RNAi was also used to identify genetic dependencies of cancer cell sensitivity to the well-known chemotherapeutics paclitaxel, doxorubicin and camptothecin[157, 172].

Although RNAi screening for a single candidate drug might have limited power, the sensitivity profile for a given candidate drug can be compared to reference datasets containing sensitivity profiles for other drugs as drugs showing a similar RNAi profile tend to share a similar mode of action. Another approach is to combine RNAi screening without drug selection and then compare the results with small molecule profiles to obtain information about mechanism of action and potential drug targets [173-175]. Finally, however, it must be noted the initial excitement of using RNAi for chemical-genetic screening has started to wane and newer methods are likely to overshadow RNAi screening in the near future[158, 162, 164, 165, 176-178].

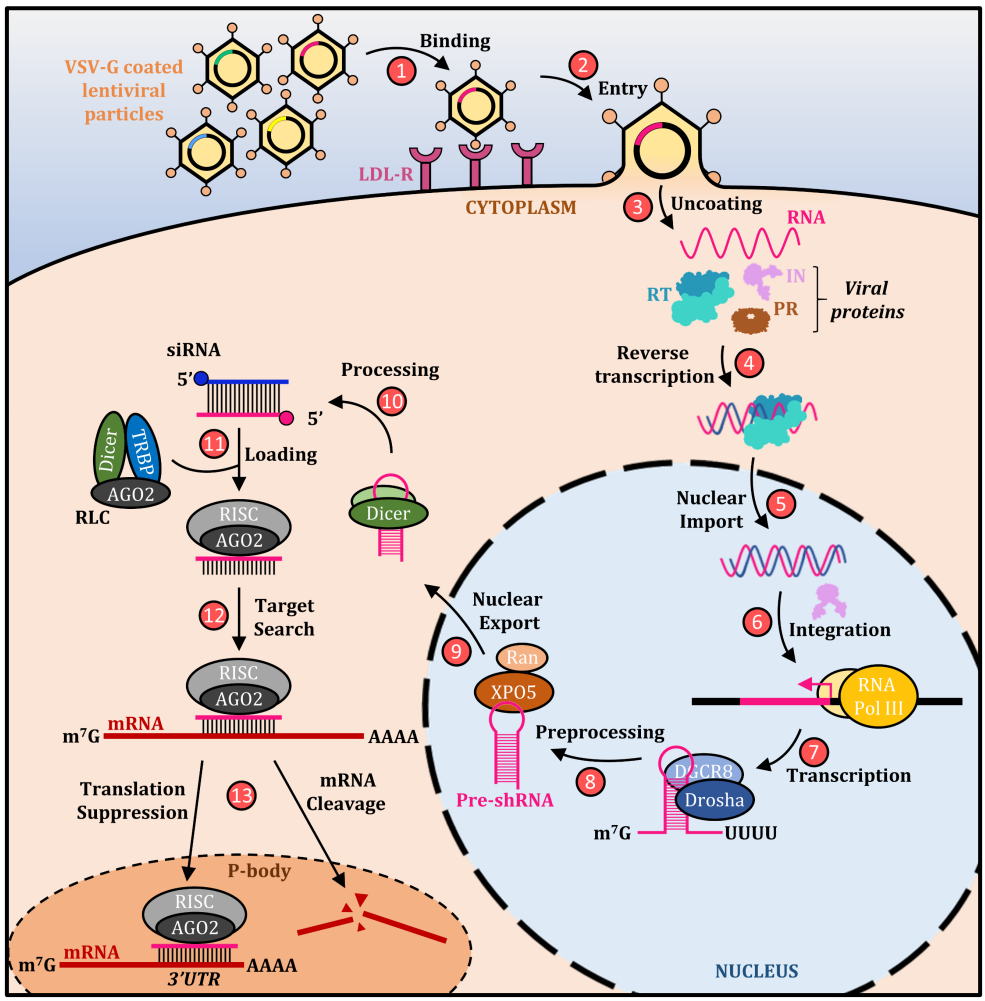


Figure 5: *lentiviral delivery of an shRNA and the mechanism of shRNA-mediated interference in a human cell*. A library of shRNA sequences is cloned and packaged into recombinant lentiviral particles pseudotyped with the vesicular stomatitis virus G (VSV-G) glycoprotein. **1.** A single lentiviral particle is entering a human cell through interactions of VSV-G with the cellular low-density lipoprotein receptor (LDL-R). **2.** The lentiviral particle fuses with the cell membrane. **3.** Lentiviral proteins, including integrase (IN), protease (PR) and reverse transcriptase (RT), and the viral RNA containing the shRNA coding sequence are released into the cell's cytoplasm. **4.** The viral RNA is transcribed into double stranded DNA by reverse transcriptase (RT). **5.** The double stranded DNA associates with viral proteins (IN, RT) and forms a pre-integration complex that is imported into the nucleus. **6.** The viral DNA is integrated into the host genome using the viral IN protein. **7.** Host RNA polymerase II or III transcribes the viral DNA into a pri-RNA molecule containing the shRNA sequence. **8.** The freshly transcribed pri-shRNA is processed into a pre-shRNA by the RNase III enzyme Drosha with support from the dsRNA binding protein DiGeorge syndrome chromosomal region (DGCR8). **9.** The pre-shRNA is bound by the Exportin-5 (XPO5)-RanGTP complex with the help of a dsRNA binding protein and exported to the cytoplasm. **10.** Once released into the cytoplasm, the pre-shRNA is further processed by the endoribonuclease Dicer into a small dsRNA fragment. **11.** The dsRNA is loaded into the multiprotein RNA-induced silencing complex (RISC) by binding to the RISC-loading complex (RLC), which consists of Dicer, Argonaute 2 (AGO2) and a transactivating RNA-binding protein (TRBP). Once loaded, one of the strands of the shRNA is degraded by the RISC while the other strand is retained. **12.** The retained sequence contains the target sequence required by the RISC to search for mRNA molecules that contain a complementary nucleic acid sequence. **13.** Once a complementary target mRNA has been bound, the RISC can interfere with its translation in two ways. I). It can stabilize the complex and suppress translation by binding to the 3' untranslated region (UTR) of the mRNA molecule. II). The Argonaute protein is activated, which will cleave the target mRNA, preventing it from being translated. Further processing and long-term inhibition of mRNA occur in so-called processing bodies (P-body), which are small but distinct foci in the cytoplasm. Of note, the targeted mRNA sequence does not have to be completely complementary to the shRNA target sequence for RISC-mediated RNA interference to occur.

1.2.3.2.2 Loss-of-function screening using insertional mutagenesis

Although RNA interference has laid the groundwork for genome-wide genetic screening in human cells, multiple issues relating to the technique have surfaced over time[158, 162, 165]. One weakness is the efficacy of RNAi for a given target gene. As RNAi relies on degradation of mRNA molecules transcribed from a particular target gene, the maximal obtainable repression for a given gene depends on the rate of mRNA transcription for that gene. In addition, different mRNA isoforms exist for many genes and due to a combination of these two observations it is difficult for genome-wide RNAi approaches to repress all target genes at similar levels. In many cases RNAi is also unable to induce a complete repression of a gene's translation and typically leads to a reduction in protein expression of 50-90%. As a more complete repression of target genes in a more consistent manner can be useful for mapping chemical-genetic interactions, researchers have turned to insertional mutagenesis.

Insertional mutagenesis strategies involve the incorporation of a genetic element into endogenous genes in the chromosomal DNA. This incorporation prevents the expression of the endogenous gene by disrupting the translation of the coding sequence due to the insertion of premature stop codons or terminator sequences together with functional splice sites, a concept known as gene trapping[38, 179-183]. Systems for gene trapping generally rely on transposon or retroviral integration elements and require large numbers of cells to generate a library of cells in which all genes have been trapped. These systems also require haploid cell models as it is almost impossible to trap multiple alleles of the same gene in diploid or multiploid mammalian cells due to their large genome size and semi-random integration of the trapping element[181, 184]. Gene traps were originally used for gene discovery and functional genetic screening focusing on only a handful of genes in model organisms such as mice and fruit flies before they were effectively applied in a genome-wide format in haploid cell lines[180, 181, 184, 185]. Gene trap mutagenesis screens have now been applied to identify genes involved in the mechanism of action and drug response of a variety of agents, although most successful examples focused on anticancer agents such as the PARP inhibitors olaparib and BMN 673, Wee1 inhibitors, the DNA-damaging agents YM155, bleomycin and doxorubicin, the BCR-ABL inhibitor imatinib, Hsp90 inhibitors and the glycolysis inhibitor 3-bromopyruvate[186-192]. Haploid insertional mutagenesis was also used to understand the mechanism of action of formaldehyde and tunicamycin[186]. In addition, as with homozygous loss-of-function profiling in yeast, these screens generally do not identify the direct target of the drug of interest and are unable to identify essential genes, as complete loss-of-function of these genes leads to cell death. Moreover, as mentioned before, they require haploid models, limiting their potential of studying multiploid models of disease.

1.2.3.2.3 Unraveling chemical-genetic interactions using gene (over)expression

Besides lowering protein expression levels to identify genes acting as suppressor or enhancer for cellular drug sensitivity, artificially increasing protein expression in the human cell can also be harnessed to probe for chemical-genetic interactions[37, 38, 146, 193]. Overexpression of specific genes can easily be obtained by transfecting or transducing cells with plasmids or lentiviruses containing the complete protein coding sequence of an endogenous gene under control of a strong promotor[194]. For example, lentiviral overexpression screens have been used to characterize genes involved in the sensitivity of cancer cells to methotrexate and other cytotoxic and DNA-damaging agents[193]. An overexpression screen was also used to map genes involved in the drug response to anaplastic lymphoma kinase (ALK) inhibitors in lung cancer[195]. This screen identified kinases that act directly downstream of ALK in addition to various modulatory proteins. Semi-random integration of promotor and enhancer elements using similar delivery systems has also been used to identify endogenous genes whose overexpression modulates cellular sensitivity to candidate drugs[196, 197]. Using these systems, modifier genes involved in sensitivity to anticancer agents such as lapatinib, erlotinib, cisplatin, paclitaxel and fludarabine have been identified. Importantly, with some exceptions, most of the

overexpression screens did not identify the direct target protein of the tested drugs and instead provide a landscape of modifier genes involved in drug mechanism of action. Besides using artificial overexpression, it is possible to correlate basal gene expression levels to drug sensitivity to predict drug mechanism of action, although this approach does not always identify the direct target protein of a drug[198]. Finally, complementation of yeast strains with human genes has also been shown to be useful for performing genetic screens and might be useful for target deconvolution[199].

1.2.3.2.4 Resistance mutations and transcriptome sequencing

Like in yeast, compelling evidence for drug-target interactions in human cells can be obtained by isolating functional mutants that confer resistance to a drug of interest in a cellular context[38, 200]. Mutations of this type typically involve mutations of only one or a few amino acids that map close to a drug's binding site on the target protein and the isolation of such mutants is regarded as the gold standard for target deconvolution[38, 200]. However, up until recently, studies relying on the isolation of point mutants for target deconvolution in a mammalian cellular context were not pursued readily due to several reasons. First, although drug-resistant human cells can be generated in a variety of ways, distinguishing relevant mutations that confer drug resistance from those that are bystander mutations is challenging as genetic methods such as backcrossing are not available. In addition, the human genome is larger and more complex than that of model organisms such as bacteria and yeast, making it costly and challenging to deconvolute mutations present in drug resistant cells. Moreover, mammalian cells, and especially cancer cells, tend to be multiploid, which can limit the discovery of resistance mutations to dominant mutations. Selection of natural drug resistance in human cells is also slow and indirect resistance mechanisms such as transcriptional reprograming and expression of off-target multi-drug resistance (MDR) protein pumps hinder the identification of point mutants[201, 202].

Regardless, different studies have shown proof-of-concept for drug target deconvolution by isolating point mutants in a mammalian context[120, 121, 203-206]. For example, multiploid HCT-116 cancer cells, which are deficient in DNA mismatch repair and have a low expression of MDR pumps, were used to derive drug resistant clones for the known anticancer agents bortezomib, ispinesib, BI 2536 and S-trityl-L-cysteine[120, 203]. Another study showed the near-haploid cancer cell line KBM-7 can be used to culture drug resistant clones against the anticancer compounds 6-thioguanine and triptolide[121]. Using transcriptome sequencing coupled to bioinformatic analysis of different resistant clones and parental cells, these studies were able to deconvolute single resistance mutations in the target gene of the used drugs, highlighting the potential of performing RNA-sequencing on drug resistant clones obtained with a haploid cancer cell line or one that has inherent mutagenic properties. Other studies report chemical mutagenesis of human cancer cells and haploid mouse embryonic stem cells accelerates the generation of drug resistance mutations against the anticancer drugs 6-thioguanine and cetuximab[204, 205]. Coupled to transcriptome sequencing and validation experiments, they were able to deconvolute dominant loss-of-function and gain-of-function resistance mutations in the target proteins. However, although these initial studies are promising, they relied on drugs for which the cellular target protein is known, and it remains to be seen whether these approaches are successful when applied to novel drugs. Moreover, if the required resistance mutations are recessive, these approaches are limited to haploid models.

1.3 CRISPR/Cas genetic engineering and screening

In the past, strategies for genetic engineering were restricted to the use of specific model systems such as bacteria, drosophila, yeast or mice. Often these strategies had low efficiencies and came with strong restrictions, requiring selectable markers and specialized tools such as transposon or recombinase systems[207-209]. Since the discovery that introduction of DNA double strand breaks allows for generating loss of function alleles and increases the efficiency for homologous recombination[210-212], many targeted endonucleases have been developed. Initially, tools for genome editing relied on chemical nucleases, meganucleases and zinc-finger nuclease proteins (ZFN), followed by the development of transcription activator like effector nucleases (TALEN)[207, 208]. Chemical nucleases are chemically engineered systems that rely on DNA-binding polymers such as triplex forming oligonucleotides (TFO) fused to a restriction enzyme or DNA-reactive agents. TFOs invade double stranded DNA through complementary. Once the DNA has been invaded, a restriction enzyme or DNA-reactive agent bound to the oligo damages the local DNA, introducing a double strand break. However, these reagents need to be redesigned for every target sequence, are generally toxic and not very specific. The meganucleases, also known as homing endonucleases, are one of the first used endonucleases and are natural occurring enzymes which recognize a unique sequence of 14-40 nucleotides. Engineering of meganucleases is possible, but not easy, as the DNA binding and nuclease domains are localized together in a single domain. As a result, engineering meganucleases requires extensive knowledge on the utilized meganuclease[207]. ZFNs and TALENs on the other hand rely on positioning small DNA binding modules in an array to facilitate DNA binding on a given target sequence. This array of DNA binding modules is then fused to a nuclease domain (e.g. FokI). This allows decoupling of the nuclease activity from DNA binding and makes engineering of these targeted nucleases easier. However, as these systems consist of a series of specific protein modules, targeting a new nucleic acid sequence will require *de novo* construction of a protein with a new array of DNA binding modules, which can especially be challenging as the DNA sequence required to encode these modules can be highly repetitive[208].

With the advent of CRISPR/Cas genome editing it is now possible to genetically manipulate virtually any DNA sequence with great ease and low cost[208, 213-215]. CRISPR/Cas genome editing utilizes a single bacterial endonuclease protein that contains nuclease and DNA/RNA binding domains. The CRISPR/Cas endonuclease requires a small single stranded RNA molecule to guide the protein to its DNA target sequence through sequence complementary. As a result, target specificity is decoupled from the protein structure and the same endonuclease can be used with different guide RNAs. This provides the system with a major advantage as it can be rapidly reprogrammed to target vastly different DNA sequences by simply redesigning the RNA guide sequence. An enormous amount of papers describing the use of CRISPR in many different cell types has been published over the last few years. This highlights how impactful, powerful and easy to use the CRISPR system really is. The only obvious limitation is that DNA cleavage by CRISPR/Cas endonucleases does not require complete complementarity between the guide RNA and DNA target sequence, which may lead to off-target DNA cleavage[216, 217]. This aspect is particularly limiting CRISPR/Cas's potential as a clinical therapeutic, although much effort is being made to reduce and control off-target cleavage[218-221]. The modular aspect of using guide RNAs to target CRISPR/Cas endonucleases shows similarities to RNAi reagents and allows for multiplexing and pooled screening[214, 222]. This turns CRISPR/Cas genome editing into a powerful genetic screening tool that can be used for drug target deconvolution and validation. Furthermore, with the continuous development of CRISPR-based platforms, CRISPR/Cas systems also provide novel approaches for genetic screening in addition to providing an alternative to almost any previously developed genetic screening approach.

1.3.1 Origins and biology of CRISPR/Cas systems

The RNA-guided CRISPR endonucleases have been derived from a prokaryotic adaptive immune system and have been optimized for use in mammalian cells[223-226]. Briefly, this natural adaptive immune system relies on the incorporation of short fragments of foreign DNA or RNA (called spacers) into clustered regularly interspaced short palindromic repeats (CRISPR) loci present in the genome of bacteria and archaea (**Fig. 6**)[227, 228]. These CRISPR loci contain small repetitive sequences flanked by different spacer sequences that provide a genetic record of foreign invaders such as viruses and pathogenic plasmids[229-232]. Once integrated into the CRISPR locus, spacers are transcribed into short CRISPR-derived guide RNAs (crRNAs) that are complementary to the genetic information making up the foreign invaders[227, 228]. The crRNAs are processed with the help of RNase and CRISPR-associated (Cas) proteins and hybridize with a separately transcribed structural RNA called the trans-activating RNA (tracrRNA). The tracrRNA-crRNA duplex is then assembled with one or more Cas effector proteins into a ribonucleoprotein complex. This CRISPR-Cas ribonucleoprotein complex then starts searching for a nucleic acid sequence complementary to the target spacer sequence of the crRNA. However, target binding of the Cas effector complex also requires a specific motif called the protospacer adjacent motif (PAM) or protospacer-flanking sequence (PFS), which is also important for the original acquisition of the spacer[227]. Once the Cas effector has bound such a motif, the target sequence is unwound and hybridizes to the crRNA spacer through complementary. Finally, once hybridization has completed, the Cas effector protein complex activates and cuts the target nucleic acid with its nuclease domains, thereby disarming the genetic element and preventing it from propagating and being expressed.

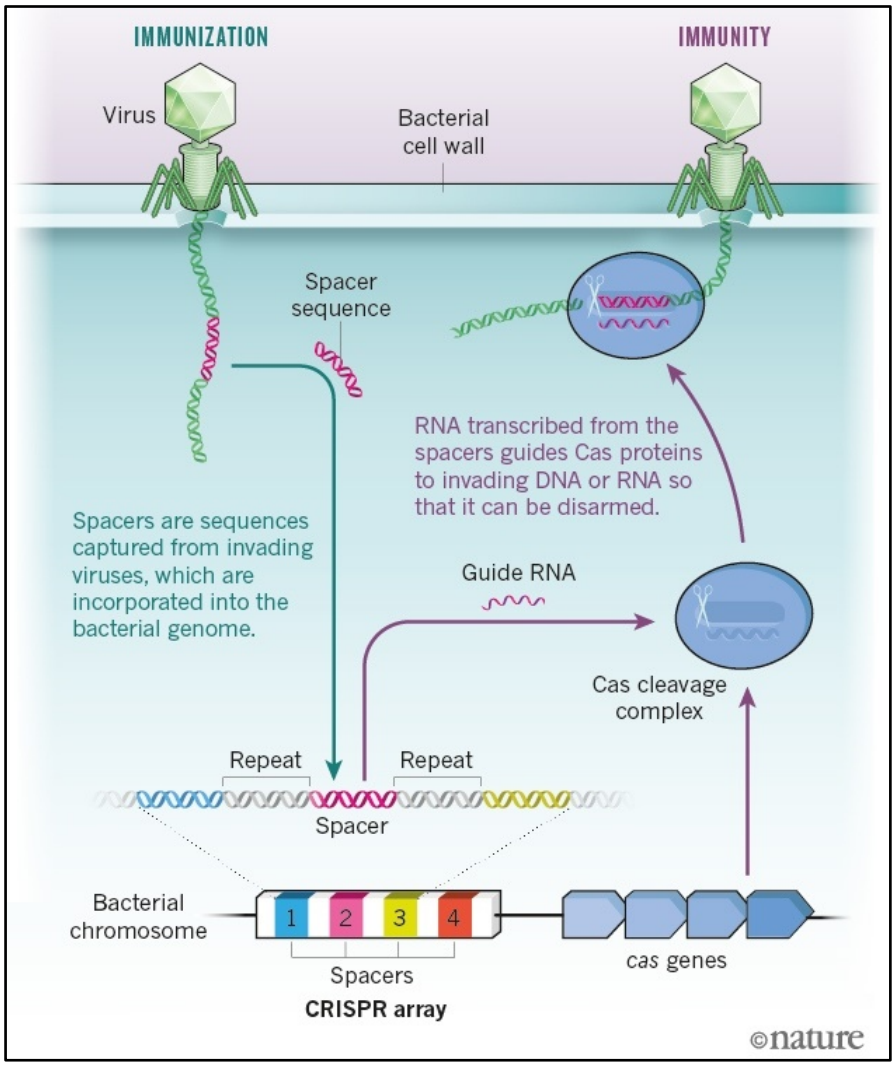


Figure 6: *schematic of the prokaryotic CRISPR/Cas immune system (adapted from Spencer/Nature).*

1.3.1.1 Classification of CRISPR/Cas systems

CRISPR/Cas adaptive immune systems show a remarkable diversity and have been suggested to evolve rapidly due to an intense arms race between viruses and prokaryotes[233-235]. The key feature of their organization is modularity as CRISPR/Cas immunity can be divided into 3 different stages[227]. The first is the acquisition and adaption of spacer sequences from foreign genetic elements. The adaptation module is largely uniform across the different CRISPR systems and typically comprises Cas1, Cas2 and Cas4 genes. After acquisition, spacer sequences and tracrRNA need to be expressed and processed to generate functional crRNA:tracrRNA duplexes. This crRNA biogenesis process differs between species but typically involves a set of proteins with RNase activity[236]. In the last step, crRNAs need to incorporate with effector proteins to form effector modules that can recognize and destroy their target elements. The composition of the CRISPR/Cas effector module shows the greatest versatility across species and forms the basis for the classification of CRISPR systems[227, 233, 235]. They are currently divided into two classes spanning 6 types that are further divided into subtypes (**Fig. 7** and **Table 1**), although it is important to note other CRISPR systems might be present in prokaryotic species that have yet to be examined[233, 237].

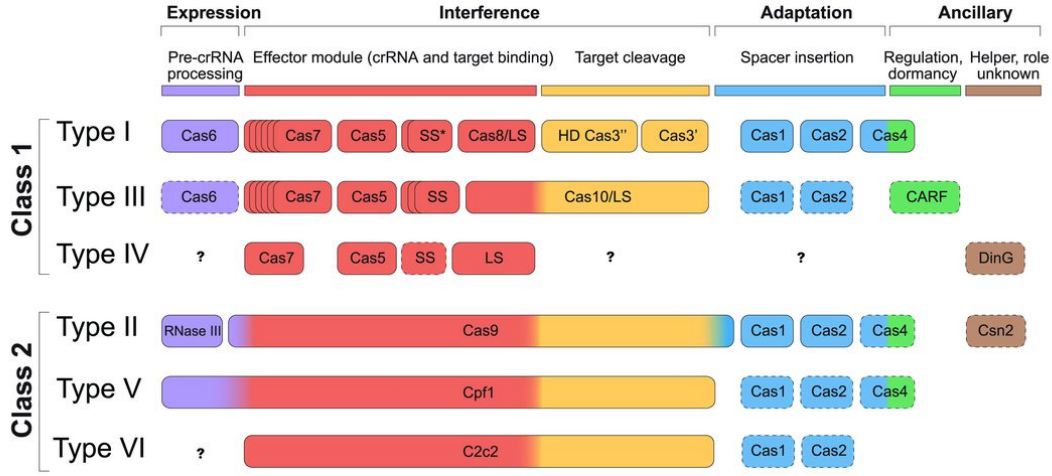


Figure 7: *overview of the modular organization of the CRISPR-Cas proteins (adapted from [235]).*

Class 1 contains the most common and diversified type I system, the type III system, present mainly in archaea, and the rare type IV system, which is remarkable by its rudimentary modules[233]. Class 1 systems are characterized by Cas effector complexes with elaborate architectures consisting of multiple proteins that generally do not require tracrRNA. On the other hand, Cas effector modules of the less common Class 2 systems are characterized by a single, large multidomain protein[233]. Class 2 systems are mainly exclusive to bacteria and include the type II system, famous for its Cas9 effector protein[221, 225], the type V system, which generally does not require tracrRNA[226], and the type VI system, which is dedicated to target RNA[238, 239]. The evolutionary origins of CRISPR systems are difficult to pinpoint but it has been suggested they might have evolved modularly by integration of multiple systems to form the acquisition and effector modules[233]. Mainly the Casposon family of transposons, transposon-encoded TnpB proteins and an ancestral RNA Recognition Motif (RRM) containing nuclease do domains have been suggested to have contributed to the formation of CRISPR-Cas systems[233].

Class	*Type*	*Occurrence*	*Subtypes*	*Main effector proteins*	*Target*	*Example*
Class 1	Type I	Common	7	Cas7, Cas5, SS, Cas3	DNA	*E. coli*
	Type III	Mostly archaea	4	Cas7, Cas5, SS, Cas10	DNA	*S. epidermidis*
	Type IV	Rare	1	Cas7, Cas5, SS, Csf1	DNA	*R. jostii*
Class 2	Type II	Mostly bacteria	3	Cas9	DNA	*S. pyogenes*
	Type V	Mostly bacteria	6	Cas12 (Cpf1)	DNA	*F. novicida*
	Type VI	Rare	3	Cas13 (c2c)	RNA	*L. shahii*

Table 1: *current classification of CRISPR/Cas systems (adapted from [233]).*

1.3.2 SpCas9-mediated genome editing

The prototype CRISPR/Cas endonuclease used for genetic engineering is the CRISPR/Cas9 endonuclease from the class 2 type 2 CRISPR system of gram-positive *Streptococcus pyogenes* bacteria[223-225]. SpCas9 is a single effector endonuclease protein consisting of 1368 amino acids. The protein comprises two major lobes termed the recognition (REC) and nuclease (NUC) lobes (**Fig. 8a** and **b**)[240-242]. The REC lobe can be subdivided into 3 recognition domains that are required to bind the crRNA-tracrRNA duplex and a bridge helix. This bridge connects the two lobes and is enriched with arginine residues essential for recognition of the crRNA:DNA duplex. The NUC lobe contains the RuvC nuclease domain, which is split up in 3 segments, the HNH nuclease domain, a topoisomerase-homology domain and the PAM interacting domain, which is essential to bind DNA. In addition, segments of both the REC and NUC lobes are required to bind DNA and facilitate formation of the crRNA:DNA duplex. Endogenously, SpCas9 requires both a tracrRNA and a crRNA to function. However, the tracrRNA and crRNA have been engineered to form a single guide RNA (sgRNA) by fusing the two RNAs together using a small loop and reducing the overall size of the tracrRNA (**Fig. 8a** right panel)[225]. This facilitates rapid repurposing of the system as only a single small RNA sequence must be cloned to target the endonuclease to a different DNA sequence. The targeting sequence of the crRNA for SpCas9 is typically 20 nucleotides long, although it is possible to shorten this sequence to 17-18 bases without losing much activity[243].

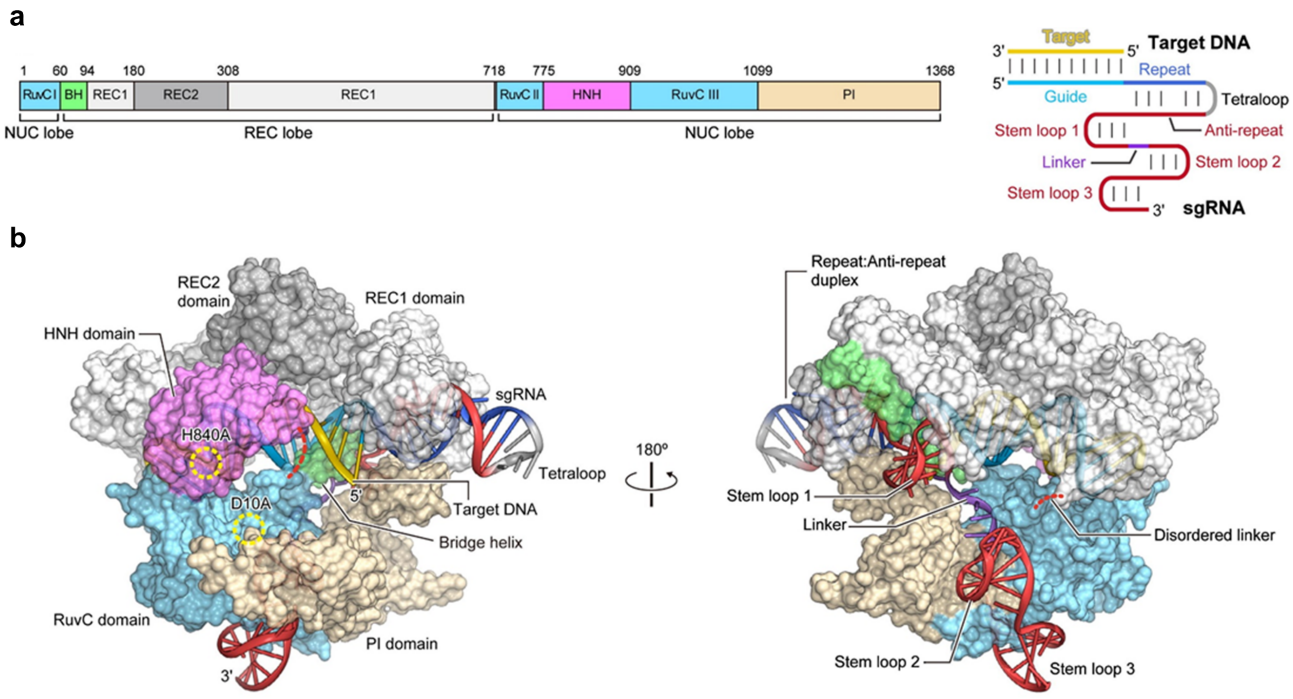


Figure 8: *structure of the Streptococcus pyogenes Cas9 endonuclease (adapted from [240]).*
a. Domain organization of SpCas9. SpCas9 is structured into two major lobes, the nuclease (NUC) and recognition (REC) lobes. The REC lobe is further divided into the REC1 and REC2 domains and the bridge helix (BH). The nuclease domain contains the HNH and RuvC nuclease domains together with the PAM interacting (PI) domain. The RuvC nuclease domain is divided into 3 subdomains. The overall structure of a single guide RNA (sgRNA) is shown on the right. In principle, the crRNA and tracrRNA are connected by a small loop (tetraloop). The target DNA is shown in yellow, the sgRNA targeting region in blue and the secondary loop structure in red.
b. Surface representation of a SpCas9-sgRNA-DNA complex. The active sites of the RuvC and HNH nuclease domains are indicated by dashed yellow circles. Single alanine mutations that can inactive the respective nuclease domains are also shown (D10A, H840A).

The SpCas9 protein requires loading of the crRNA-tracrRNA (or sgRNA) to show endonuclease activity. Once the required RNA has bound SpCas9, the protein will actively search the environment for DNA sequences containing the NGG protospacer adjacent motif. Once it has found such a PAM, SpCas9 unwinds the 5' proximal DNA to allow the crRNA target sequence to hybridize with the DNA that is complementary to the string on which the PAM is located. Thus, the non-target strand is the DNA sequence that contains the NGG PAM. The recognition of the target DNA sequence is especially dependent on the hybridization of the first 10-12 nucleotides proximal to the PAM site, which is termed the seed region (**Fig. 9**)[224, 225, 243, 244]. The more PAM-distant nucleotides of the sgRNA tolerate mismatches and are not as essential for target recognition. Once hybridization has completed, the HNH domain, which is flexible and relatively unstructured in the unbound state, is locked into position on the target strand[221, 241, 242]. This conformational lock on the HNH domain also activates and locks the RuvC domain on the non-target strand. Both strands of the hybridized DNA are then cut concurrently by the activated RuvC and HNH domains[221].

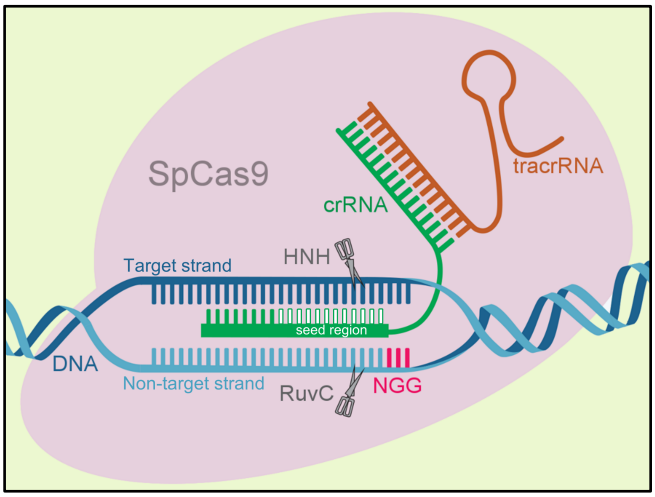


Figure 9: *schematic of SpCas9 binding to target DNA (adapted from Integrated DNA technologies).* The SpCas9 endonuclease is shown as a pink silhouette. It has bound the CRISPR RNA (crRNA, green) and tracer RNA (tracrRNA, orange) duplex. Once bound to the duplex, SpCas9 searches the environment for dsDNA containing a PAM sequence (NGG, purple). Once a PAM is recognized, the DNA is unwound. The crRNA then hybridizes to the DNA and forms an R-loop. When hybridization finishes, the complex stabilizes and the HNH and RuvC endonuclease domains of SpCas9 will lock on and cut the DNA on both strands to generate a double strand break in the target DNA.

1.3.2.1 Repair of DNA double strand breaks

The majority of genome editing approaches rely on the introduction of a DNA double strand break (DSB) in the DNA of the organism to be modified. These DSBs are repaired by the organism using its own DNA repair pathways and this machinery is used for genome editing to introduce the genetic modification of interest[208, 245, 246]. Typically, two DNA repair pathways are distinguished (**Fig. 10**)[247]. Non-homologous end-joining (NHEJ) is the first of these pathways and is active in nearly all cell types. In simple terms, NHEJ rejoins the ends of DNA fragments that were generated by the double strand break[245]. This process is inaccurate and often leads to the introduction of deletions, insertions and substitutions at the break site. Homology-directed repair (HDR) on the other hand relies on copying the genetic information from a homologous template to the site of the DSB and is very precise, but generally inefficient and only active during mitosis[245]. It is perhaps not surprising that both pathways rely on a complex set of host proteins that are devoted to repairing DNA damage as DNA double strand breaks are dangerous lesions that, if not properly repaired, can lead to cell death or disease.

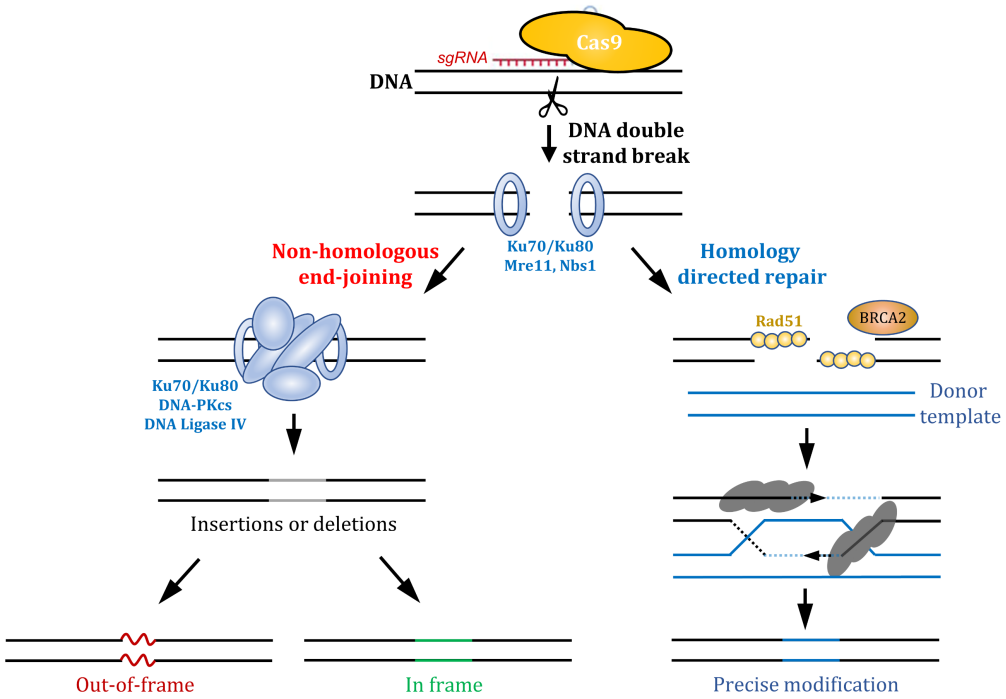


Figure 10: *overview of DNA repair outcomes of a CRISPR/Cas9-induced DNA double strand break.* When the CRISPR/Cas9 endonuclease has recognized its target site, it will cleave both strands of the target DNA. This leads to the formation of a DNA double strand which is almost immediately recognized and bound by the Ku70/Ku80 and Mre11/Nbs1 DNA repair proteins. The DNA can now be repaired in two ways. The first is non-homologous end joining (NHEJ) in which additional DNA repair proteins are recruited to ligate both ends back together. However, before ligation, exonucleases and polymerases can remove or add a small number of bases. This process thus becomes error-prone and often leads to the formation of insertions and deletions, which, if localized to a protein coding sequence, can lead to out-of-frame or in frame mutations. The other DNA repair pathway is homology directed repair (HDR). This process is generally only active during mitosis and requires the DNA repair proteins RPA, Rad51 and BRCA2. The DNA is first resected by Mre11, and a template DNA sequence that is homologous to the sequence surrounding the DSB is searched to function as repair template. Once a template has been found, strand invasion of the template DNA occurs, and the sequence of the repair template is copied over by polymerases into the DSB. Under normal cellular conditions, the donor repair template typically consists of the sister chromosome. However, a repair template can also be provided artificially in order to manipulate the damaged DNA as desired. The outer arms of the donor DNA fragment need to be homologous to the DNA sequence surrounding the DSB to facilitate HDR.

1.3.3 Nuclease-dead variants of CRISPR/Cas endonucleases

The CRISPR toolbox has been expanded further with the development of catalytically dead CRISPR/Cas variants. For example, nickase mutants have been developed in which either the active site of the RuvC or HNH-like nuclease domain of SpCas9 was inactivated by mutation (D10A, H840A respectively)[208, 225]. These mutants can be used to nick the target DNA sequence by introducing a cut in only one strand. These nicks still allow for the induction of homology-directed repair but lower the occurrence of NHEJ and can be used to increase the specificity of genome engineering with CRISPR[208, 248]. A catalytically dead SpCas9 (dCas9) can be generated by combining inactivating mutations in both the RuvC and HNH nuclease domains[213]. However, dead SpCas9 can still bind sgRNAs and search for PAM sequences to find target DNA sequences. This effectively turns dCas9 into a highly specific DNA-binding platform, which allows for recruitment of functional elements to DNA target sites to manipulate DNA without inducing mutations[213, 249]. For example, specific DNA sequences can be visualized by recruiting fluorescent molecules to DNA-bound dCas9. It is also possible to recruit epigenetic factors to modulate gene expression or recruit deaminases to allow the introduction of point mutants without inducing strand breaks[213, 249-252].

1.3.4 Genetic screening with CRISPR/Cas and its use for target deconvolution

As CRISPR/Cas endonucleases are guided to DNA target sequences by easily designed small RNA molecules, genetic engineering with CRISPR/Cas is well suited for large scale genetic screening. Generally, expression constructs for sgRNAs in CRISPR-based methods are stably inserted into the genome by lentiviral transduction. Integrated sgRNAs can then directly serve as DNA barcode as part of their nucleotide sequence is identical to a part of the genetic element it is targeting. Many different platforms for CRISPR/Cas genetic screening have been developed for use in mammalian cells. Here we aim to provide an overview of these approaches and how they have been used for drug target validation and deconvolution.

1.3.4.1 CRISPR loss of function screening

CRISPR/Cas-induced NHEJ repair often leads to the generation of loss-of-function (or knockout) alleles due to frameshift mutations[214, 222, 253]. This principle can be applied in a chromosome-wide fashion to probe almost every gene or functional genetic element as CRISPR allows to target genetic information at its source, the DNA. Up until now, gene loss-of-function screening in mammalian systems had been limited to haploid model systems. This requirement was set by the only available alternative: gene traps[179-181, 185, 254-256]. Gene trap systems cannot be effectively applied to diploid or multiploid systems as they rely on lentiviral or transposon elements that randomly integrate into the genome, making it almost impossible to obtain knockouts of the same genetic element on different alleles. Fortunately, loss-of-function/knockout screening with CRISPR works in diploid and multiploid model systems, although it is likely that its efficiency drops with increasing multiploidy[214, 222, 257, 258]. By no doubt CRISPR genome-wide loss of function screening is becoming one of the strongest genetic screening approaches now available, in part due to its robustness, ease of use and simplicity. However, it does come with its own issues. For example, repeated genetic elements present in high copy tend to score high in dropout screens when targeted by CRISPR/Cas endonucleases, even when these elements do not encode any genetic element involved in the probed phenotype[259]. This is because excessive induction of double strand breaks in the DNA by CRISPR/Cas cutting is toxic to the cell. However, an algorithm to account for this observation has already been developed[260]. Furthermore, early genome-wide CRISPR sgRNA libraries were not well-optimized for knockout generation, but improvements are actively being made and currently available libraries are quite optimized[261, 262]. Genome wide CRISPR knockout has been compared to gene traps in haploid cell systems and has been shown to be just as or potentially more effective as gene traps in these model systems[185, 263, 264]. In addition, genome wide CRISPR loss-of-function screening has been compared to RNAi with the conclusion that CRISPR is as potent as or better than RNAi, although both approaches can complement each other to obtain high quality datasets[176, 265-267]. Besides these comparative studies, many studies have used CRISPR knockout screening to answer biological questions. For example, CRISPR knockout screening has been used to identify essential genes in many studies examining both normal cells and different cancer cell types[185, 214, 222, 259, 268-275]. Besides these studies, CRISPR knockout screening has also been adopted widely to probe for genes modulating drug or toxin resistance[214, 222, 257, 265, 276-283]. However, these studies typically did not identify the target of a candidate drug, unless the target gene was non-essential. Indeed, CRISPR knockout screens for drug sensitivity do not perform well for essential genes, as loss of such genes is lethal.

1.3.4.2 CRISPR interference and activation

RNA interference ushered the start of genetic screening in human cells by providing a programmable system that can reduce target gene expression through interference with its messenger RNA on a large scale[160, 161, 164]. Now a few years ago, catalytically dead SpCas9 (dCas9) fused to transcriptional repressors such as KRAB and p300 has been shown to effectively inhibit mRNA expression at the level of the promotor, a process termed CRISPR

interference (CRISPRi)[284-286]. Although targeting of the dCas9-repressor requires knowledge on transcription start sites and nucleosome occupancy[287], genome-wide CRISPRi libraries with reasonable efficiency have been developed[253, 286, 288]. Moreover, as an exogenous system, CRISPRi, unlike RNAi, does not require endogenous machinery to process its reagents. CRISPRi also interacts directly with the DNA, avoiding problems with abundant or alternatively spliced mRNA. Besides being able to reduce expression levels, CRISPR systems have also been developed to provide an alternative to overexpression screens. The fusion of dCas9 to transcriptional activators such as VP64 and p65, termed CRISPR activation (CRISPRa), allows for the activation of gene transcription by targeting endogenous transcription start sites in the DNA[213, 286, 289]. Genome-wide libraries for use with CRISPRa have been developed, although there is still room for improvement[286, 290]. Nevertheless, both CRISPRa and CRISPRi can modulate gene expression levels up to 1000-fold and allow to probe essential genes. In addition, genetic perturbations with CRISPRi and CRISPRa are reversible as the targeted genetic element is not permanently altered due to repair of DNA double strand breaks. However, in contrast to RNAi and classical overexpression screening, CRISPRi and CRISPRa modulate their target elements by binding directly on the chromosome. This localization could influence other unrelated genetic elements near the target site. Overexpression screens with plasmid or lentiviral open-reading-frames require genetic elements of various lengths as each expressed gene is unique. However, CRISPRa utilizes sgRNA vectors that are all the same size and almost identical. The only difference is the variable 20-bp sgRNA spacer sequence for targeting endogenous genes. Because of these differences and advantages, CRISPRi and CRISPRa genetic screening has been applied to study chemical-genetic interactions. One of the first screens combined CRISPRa and CRISPRi in a small-scale screen to probe genes involved in ricin sensitivity. Using this data, they generated genome-wide CRISPRa and CRISPRi libraries to screen for factors modulating sensitivity of human cells to cholera-diphtheria toxin[288]. Small scale CRISPRa and CRISPRi experiments were also used to validate the AAA ATPase p97 as the target of CB-5083[291]. Another small-scale CRISPRi screen in the yeast *Saccharomyces cerevisiae* characterized the mechanism of action of various small molecules[292]. While a genome-wide CRISPRa library identified EGFR, BCAR3, a list of G-protein-coupled and ITG receptors as modulators of the response to the BRAF inhibitor PLX-4720[289]. A CRISPRa screen identified B4GALNT2 as a restricting receptor preventing infection with many avian influenza viruses[293]. Combined genome-wide CRISPRa and CRISPRi screening have also been used to identify the cellular target of the anticancer agent rigosertib and to validate the target of the BRAF inhibitor vemurafenib[290, 294]. CRISPRi screening has also been used to map genetic interactions of human factors involved in chromatin-regulation[295] and to map and identify essential genes in bacteria and human cells[275, 296, 297]. Both CRISPRi and CRISPRa have also been used to study the non-coding genome[298-300]. Finally, CRISPRi has been applied in a single cell fashion to dissect the unfolded protein response[301].

1.3.4.3 Utilizing CRISPR-targeted base editors to generate point mutants

The generation of point mutations across the genome to obtain novel gain-of-function mutations might be the most difficult, but arguably also the most powerful, genetic approach[38, 302]. This approach has the potential to interrogate complete genomes at base-pair resolution and has been applied to lower eukaryotes for a long time. Just recently these approaches have been applied to mammalian cells in an unbiased manner to identify the cellular target of anticancer agents[204, 205]. Current methods rely on the use of chemical mutagens or UV irradiation to induce random mutations across the genome. Although powerful, the efficiency of generating gain-of-function mutations in desired genes using these methods is low due to the random nature of mutagenesis coupled to the small fraction of functional genetic elements in the entire genome. Moreover, obtaining the same mutation in multiple alleles is unlikely. In addition, multiple mutations in a single genome are often produced, hindering the

deconvolution of relevant mutations. Although the continued development of next generation sequencing technologies and analysis tools will probably negate this limitation over time, the recent development of CRISPR/Cas base-editing tools highlight CRISPR/Cas can also be used to generate point mutants or gain-of-function alleles on a medium, semi-random scale[250, 303-305]. CRISPR/Cas base editors can be divided into two classes, precise editing tools generating single predefined point mutations in a small window and hyperactive variants that tend to randomly mutate a few bases within a window of ten to a hundred bases around the sgRNA binding site. These hyperactive base-editors rely on the recruitment of variants of the activation-induced cytidine deaminase (AID), a protein involved in somatic hypermutation in the immune system, to dCas9 to create a targetable mutagenic platform. These CRISPR/Cas AID-coupled base-editors have been applied in a pooled format to scan the target gene of the anticancer agents bortezomib and imatinib for drug resistance mutations in human cancer cells[250, 303, 304]. Although these base-editors have only been applied to mutagenize single genes, they should be applicable to mutagenesis scanning of multiple genes in a single experiment. Further improvements to these systems might provide researchers with tools to perform CRISPR-driven evolution in human cells on a large scale, although probably still nowhere near the size of the whole genome. Interestingly, hyperactive base editors can also be used as multiplex systems, allowing rational mutagenesis of multiple sites in a single cell[303].

1.3.5 Future of CRISPR/Cas genetic engineering

Besides its obvious use for functional genetic screening, the CRISPR/Cas toolbox also facilitates genetic engineering experiments on a small scale. This includes the generation of mutant cell lines, the generation of animal models of disease, manipulating primary cells and performing quick genetic validation experiments[227, 306]. This aspect of CRISPR/Cas genome editing is rapidly being adopted by biomedical and biological labs across the world with CRISPR/Cas genome editing evolving to become a standard in any lab. Various institutes are developing CRISPR for therapeutic use to correct genetic diseases, although off-target effects are still a major problem[307, 308]. Besides generating more specific CRISPR/Cas endonucleases with lower off-target effects, other aspects to address are the development of additional CRISPR/Cas endonucleases targeting a diverse repertoire of PAMs, obtaining a better understanding of DNA double strand break repair, improving the occurrence of HDR, and improving methods to characterize off-target effects. In addition, *in vivo* applications of CRISPR and genome editing come with major regulatory and ethical issues and challenges, especially concerning germline editing. Thus, a framework on the *in vivo* use of CRISPR-Cas9 technology is urgently needed[309, 310]. Another ethical debate gaining ground is the concept of gene drives[311]. Gene drives are small genetic elements that are inherited through a population in a super-Mendelian manner, where the genetic element is copying itself onto genomes that previously lacked it. After a few generations, this can lead to populations where almost every individual contains the genetic drive element. CRISPR/Cas gene drives have been used to generate populations of infertile or malaria resistant mosquitos in the lab[312, 313]. CRISPR/Cas genome editing also holds great potential for altering traits in plants and livestock. For example, sheep and cattle have been edited using TALENs to carry more muscle mass[314]. Multiple other examples exist, but with the continued development of genome editing tools the possibilities continue to expand[315].

It remains to be seen whether CRISPR/Cas systems will become the new norm for genetic screening, but the signs bode well[177, 178, 266, 316]. In addition, new territory will probably be continued to be explored. The fusion of catalytically dead Cas9 to epigenetic regulators such as DNA methyltransferases and histone deacetylases has already been performed[213, 317, 318]. Besides these epigenetic regulators, fusion of dCas9 to structural domains to explore chromosome structure and compaction are also expected[319]. Finally, CRISPR/Cas systems also provide great tools for synthetic biologists and we will likely see more adaptations of CRISPR/Cas to generate sophisticated genetic circuits and systems[320-322].

1.4 Exportin-1-mediated nuclear export

1.4.1 Introduction to nuclear-cytoplasmic transport

The nucleus is the defining characteristic of the eukaryotic cell. It protects the genome and forms a tightly controlled environment that allows the cell to regulate cell division and transcription[323]. The nucleus is surrounded by a nuclear envelope or nuclear membrane that consists of a double lipid bilayer membrane which functions as a barrier for the nuclear exit and entry of molecules[324]. Nuclear pore complexes (NPC) are embedded into the nuclear envelope and form selective, large multi-protein channels that function as gatekeepers for the trafficking of molecules between the nucleus and cytoplasm[325, 326]. Small molecules such as ATP, metabolic products and light proteins can travel through the NPC by passive diffusion. However, macromolecules including large proteins, RNA and protein complexes cannot cross the NPC through passive diffusion. Instead, these large molecules are transported across the NPC in an energy- and receptor-dependent manner. Much of this active transport is carried out by the karyopherin-β family of nuclear transport receptors[324, 327]. The human family consists of 18 proteins that are commonly subdivided into the β-importins and the exportins[327-329]. These proteins consist of ~20 consecutive pairs of antiparallel α-helices, which are arranged to form super-helical or ring-shaped proteins. Each of these helical pairs spans about 40 amino acids in which both helices are connected by a small loop. As their names imply, β-importins facilitate the transport of macromolecules from the cytoplasm to the nucleus, while the exportins facilitate the transport of cargo molecules from the nucleus to the cytoplasm. Although most karyopherin-β family proteins can function alone, the main β-importin, known as karyopherin subunit beta 1 (KPNB1), requires adaptor proteins, known as the importin-α (karyopherin-α) family of proteins, to facilitate nuclear import of a subset of its cargo molecules[330].

1.4.2 The nuclear-cytoplasmic transport cycle

Importins and exportins recognize specific signal sequences present in their cargo proteins to facilitate transport across the nuclear envelope[327, 331, 332]. These signal sequences are separated into nuclear localization signals (NLS), which are recognized by importins, and nuclear export signals (NES), which are recognized by exportins. In addition to recognition of cargo molecules through these signal sequences, karyopherin-mediated nuclear transport also requires energy in the form of a small Ras related GTPase called Ran. GDP-bound and GTP-bound forms of Ran are present in the nucleus and the cytoplasm, with a high concentration of RanGTP in the nucleus and RanGDP in the cytoplasm. This RanGDP-GTP gradient across the nuclear membrane controls the assembly and disassembly of cargo-karyopherin complexes and provides directionality to karyopherin-mediated nuclear transport. In the nucleus, NES-containing cargo proteins bind to exportins in the presence of RanGTP to start the nuclear export process (**Fig. 11**)[333]. Then, when the exportin-RanGTP-cargo complex crosses the NPC to the cytoplasm, RanGTP is displaced from the complex by Ran binding proteins (RanBP) and activated by Ran GTPase activating proteins (RanGAP). This triggers dissociation of the cargo molecule from the exportin complex and Ran-mediated hydrolysis of GTP to GDP, disrupting the affinity of Ran to the exportin. The opposite happens for nuclear import. Importins only bind to NLS-containing cargo proteins in the cytoplasm in the absence of RanGTP. Once the importin-cargo complex has translocated to the nucleus, the cargo-importin interaction is disrupted by binding of the importin to RanGTP, resulting in the release of the cargo protein in the nucleus. Importantly, the amphipathic α-helical repeats making up the β-karyopherin proteins expose hydrophobic patches on their outer surface. These patches allow β-karyopherins to cross the NPC through interactions with stretches of Phe-Gly (FG) repeats present on NPC proteins. These FG repeats are intrinsically disordered and, together with bound karyopherins, form the permeability barrier in the center of the NPC[325, 326, 334]. They prevent free passage of macromolecules, while they promote the karyopherin-mediated transport of cargo molecules through hydrophobic interactions.

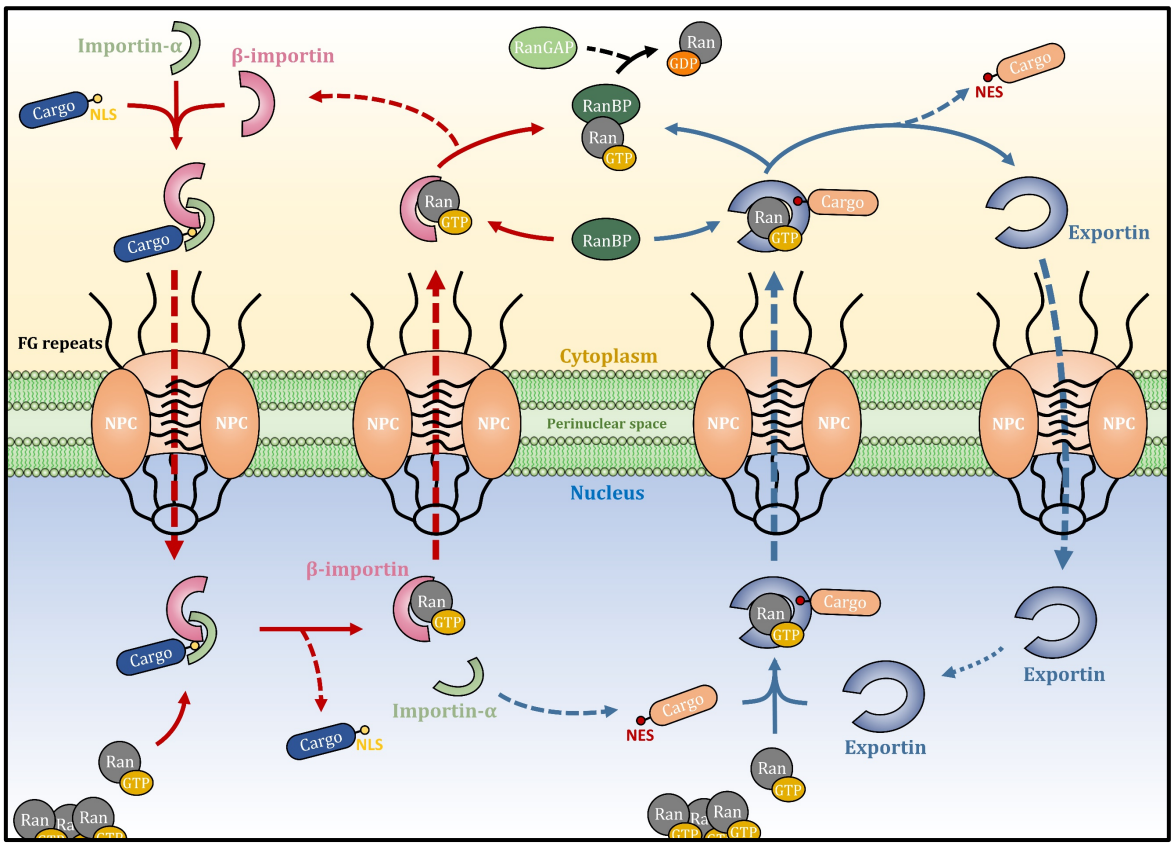


Figure 11: *overview of karyopherin-mediated nuclear-cytoplasmic transport.*
NPC: nuclear pore complex. **NLS:** nuclear localization signal. **NES:** nuclear export signal.

1.4.3 Exportin-1

Exportin-1 (XPO1), also known as chromosome region maintenance 1 (CRM1), is the major nuclear export factor that is responsible for the nuclear export of more than hundreds[335-337], if not up to a thousand[338, 339], cargo proteins from the nucleus to the cytoplasm[340-344]. Its cargo includes many tumor suppressor factors, growth regulatory proteins, redox-related proteins, ribosomal proteins and other RNA binding proteins. As a result, XPO1 is one of the best-studied karyopherins and the XPO1-mediated nuclear export has been implicated in various diseases such as cancer, inflammation, heart failure, viral infection, neurodegeneration and wound healing[327, 345-356]. Overexpression of XPO1 has been reported in different cancer types and is associated with poor prognosis and drug resistance[327, 357-363]. Moreover, missense mutations of the glutamic acid residue 571 of XPO1 have been linked to cancer, although the molecular consequences of these mutations are still unknown[364-368]. XPO1 has also been reported as a specific cancer dependency in a subset of non-small lung cancer and sarcoma[369, 370]. These findings highlight the potential of modulating XPO1 function for the treatment of disease and, not surprisingly, various inhibitors of XPO1 function have been described.

1.4.3.1 Structure of XPO1

As for all karyopherin-β members, XPO1 consists of a series of consecutive pairs of antiparallel α-helices (named the A- and B-helix) connected to each other by a small loop (**Fig. 12a**). These repeats are also known as HEAT repeats, named after the presence of these helix repeats in the "**H**untingtin", "**E**longation Factor 3", "Protein Phosphatase 2**A**" and "**T**arget of Rapamycin" proteins[371]. In the case of XPO1, 21 of these repeats are arranged to form a ring-shaped protein resembling a flexible open donut (**Fig. 12b**)[332, 372, 373]. The outer surface of the HEAT repeats is made up by the A-helices that allow for interaction with cargo and FG repeats on the NPC. The inner surface of the protein is formed by the B-helices and plays a role in the binding of RanGTP and cargo. Once RanGTP is bound, the open-donut shape of the protein becomes more rigid and

appears to close (**Fig. 12c**). The cargo-binding site consists of multiple α-helices that make up a NES-binding cleft rich with hydrophobic patches at HEAT repeats 11 and 12 (**Fig. 12b** and **c**). Remarkably, the B-helix of the C-terminal HEAT repeat 21 folds onto the XPO1 structure in the free, unbound state (**Fig. 12b**)[374]. This folded C-terminal tail interacts with the base of the acidic loop of HEAT repeat 9, which pushes against the NES-binding cleft (**Fig. 12b**). This interaction forces the NES-binding cleft into a rigid and closed state, which prevents the interaction of XPO1 with its cargo. Binding of RanGTP to XPO1's inner core pushes this C-terminal tail outwards to prevent steric clashing. This conformational change frees up the acidic loop and the NES-binding cleft (**Fig. 12c**)[374]. This facilitates a conformation in which the NES-binding cleft opens so that it can accessed by the nuclear export signal sequence present on XPO1's cargo proteins.

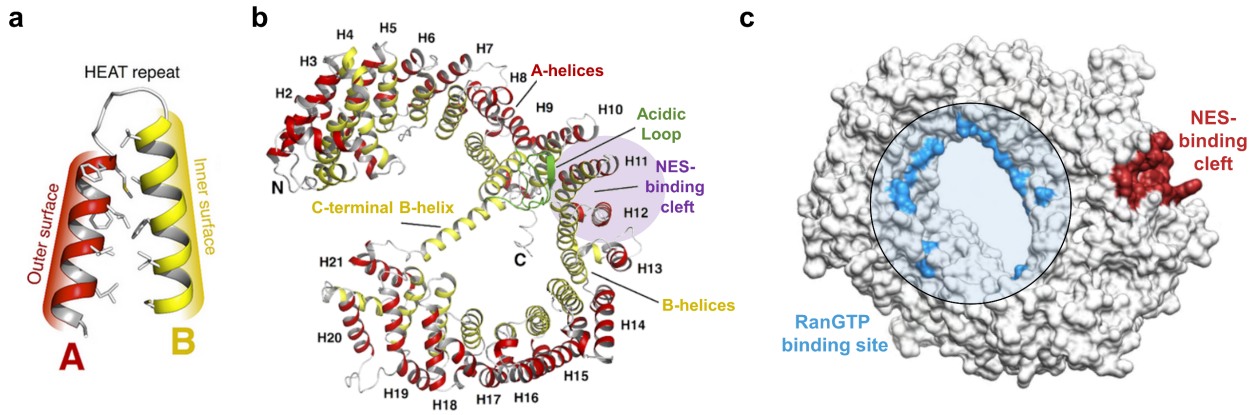


Figure 12: *structure of the karyopherin-β nuclear export receptor XPO1 (adapted from[372-374]).*
a. Structural organization of a HEAT repeat. The HEAT repeat is formed by two antiparallel α-helices, called A and B, connected by a small loop. A single HEAT repeat typically encompasses 40-50 amino acids. The HEAT repeat is further stabilized by hydrophobic interactions between the A- and B-helices.
b. Ribbon diagram of the structure of XPO1. The HEAT repeats are shown with the A- and B-helices colored in red and yellow respectively. The A-helices form the outside of the protein while the B-helices form the inner core. The NES-binding cleft is formed by the HEAT repeats 11 and 12 (purple) and interacts with the acidic loop of HEAT repeat 9 (green). The C-terminal tail encompassing the B-helix of HEAT repeat 21 is shown interacting with the acidic loop.
c. The surface of RanGTP-bound XPO1 is shown (RanGTP not shown for clarity). When XPO1 binds RanGTP, the open-donut structure becomes more rigid, closes and pushes out the C-terminal tail from the inner core of the structure. The RanGTP binding surface is highlighted in blue while the surface of the NES-binding cleft is highlighted in red.

1.4.3.2 Cargo recognition by XPO1

In contrast to other exportins, exportin-1 interacts with its cargo proteins on the outer surface of its structure. Many different cargo proteins can bind XPO1 and interact with XPO1 through an exposed nuclear export signal (NES) present on the cargo protein. The NES typically encompasses 8-15 amino acids and docks onto 5 hydrophobic patches located in the NES-binding cleft of XPO1 (**Fig. 13**). NES binding to XPO1 generally requires preformation of the RanGTP-XPO1 intermediate, although some cargo proteins can bind XPO1 without requiring RanGTP[374]. The preformation of RanGTP-XPO1 as a requirement for cargo-binding also provides XPO1-mediated nuclear export its directionality. Moreover, the nuclear RanBP3 protein has been shown to stabilize the RanGTP-bound conformation of XPO1, further supporting the formation of cargo-RanGTP-XPO1 complexes in the nucleus[375].

The human immunodeficiency virus-1 (HIV-1) protein Rev was the first described cargo protein of XPO1 and contains a leucine-rich nuclear export signal sequence which docks onto the XPO1 structure (**Fig. 13** top right panel)[341, 343, 344, 351, 355, 376]. Besides this viral protein, many other XPO1 cargo proteins have now been identified. Generally, they all contain a nuclear export signal containing four to five spaced hydrophobic amino acids (Φ-residues)[373]. The overall

structural properties of the NES can vary between cargo proteins, requiring different arrangements of the Φ-residues to compensate for differences in spacing[332, 338]. For example, the binding of the nuclear export signal of the XPO1 cargo protein kinase A inhibitor (PKI) is different from the binding mode of Rev's NES[373]. The PKI NES adopts a more α-helical form and the hydrophobic amino acids interacting with XPO1 are spaced by more amino acids (**Fig. 13** compare right panels). Besides binding of the NES sequence, XPO1 can also interact with some cargo proteins, such as Snurportin-1 and Rev, through other sites on its outer surface[332, 374, 377]. These interactions are believed to further stabilize the cargo-XPO1 interaction and enhance binding of some cargo proteins. Together with the differences in structural NES properties between different cargo proteins, it becomes obvious that cargos bind XPO1 with different affinities. Indeed, dissociation rates between different cargo proteins can range between low nanomolar to high micromolar values[374]. These differences provide a dynamic layer of regulation to nuclear transport. Controlling the expression or availability of a cargo molecule and its NES allows fine-tuning of the nuclear export efficiency for this cargo and also for other NES-containing cargo proteins through spatial-temporal competition and posttranslational modifications[378, 379].

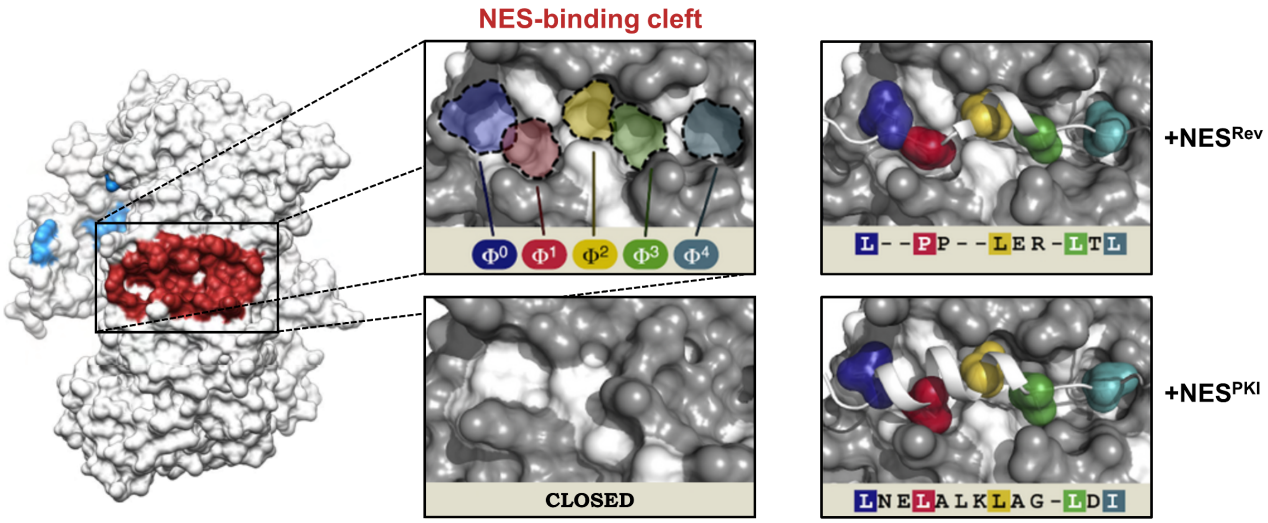


Figure 13: *close-up view of NES-binding to the XPO1 NES-binding cleft (adapted from:[373, 374]).*
Left panel: The surface of RanGTP-bound XPO1 is shown (RanGTP not shown for clarity). The image is a turned view of the crystal structure shown in **figure 12c**.
Top middle panel: the 5 major hydrophobic patches in the nuclear export signal (NES)-binding cleft of XPO1 are shown. The image was obtained from the crystal structure of XPO1 bound by RanGTP and its cargo protein Snurportin-1 (Snurportin-1 not shown for clarity).
Bottom middle panel: the closed NES-binding cleft of XPO1 is shown. The image was obtained from the crystal structure of free XPO1.
Top right panel: the interaction between the nuclear export signal of the viral HIV-1 protein Rev with XPO1 is shown. Key hydrophobic residues (Φ-residues highlighted in the amino acid sequence) of the nuclear export signal from Rev are interacting with the hydrophobic patches in the XPO1 NES-binding cleft. The image was obtained from the crystal structure of XPO1 bound to RanGTP and Rev.
Bottom right panel: the interaction of the protein kinase A inhibitor (PKI) nuclear export signal with XPO1 is shown. The Φ-residues of the nuclear export signal from PKI are highlighted in the amino acid sequence. The image was obtained from the crystal structure of XPO1 bound by RanGTP and PKI.

1.4.3.3 Small molecule inhibitors of XPO1

XPO1 was first described in yeast as chromosomal region maintenance 1 (CRM1) due to the discovery of cold-sensitive mutations in an essential nuclear protein that was important for higher order chromosomal structure[380]. Soon, a *CRM1* homolog was discovered in human cells and was shown to facilitate the nuclear export of HIV Rev, PKI and U snRNAs[340-344]. In addition, it was shown that CRM1 can interact with the FG repeats of the nuclear pore complex and it was suggested to rename CRM1 to Exportin-1 (XPO1)[341, 344, 381, 382]. Around the same time, a gene that conferred resistance to leptomycin A and B, toxic antifungal agents produced by *Streptomyces* bacteria, was reported in a *Schizosaccharomyces pombe* yeast strain[383-385]. This gene was *CRM1*, which suggested that these natural products act by binding to XPO1/CRM1. Further analysis confirmed that these natural products indeed bind to XPO1[340, 386-388]. Other bacterial agents binding to XPO1 have been discovered since then and include anguinomycin A, B, C and D and ratjadone A and C. All these bacterial inhibitors, including the leptomycins, are potent XPO1 inhibitors characterized by their structure, which consists of a polyketide chain coupled to a lactone ring[389]. Various other natural inhibitors of XPO1 have now been reported[389, 390]. Interestingly, all these XPO1 inhibitors bind XPO1 through a covalent Michael addition on the cysteine$_{528}$ residue located in the hydrophobic cargo binding cleft of XPO1 (**Fig. 14** left and middle panel)[386, 387]. This binding has been shown to abolish the interaction of cargo proteins with XPO1, thereby blocking XPO1-mediated nuclear export of cargo. Although most of these XPO1 inhibitors show anticancer activity *in vitro*[391-393], leptomycin B was the only one to enter phase 1 trials for the treatment of cancer[394]. However, leptomycin B showed poor clinical benefits due to strong dosage-limiting toxicities[394]. Besides the discovery of these natural products, several efforts have led to the discovery of synthetic small molecule inhibitors of XPO1[389, 390, 395]. For example, analysis of a series of small molecule inhibitors of HIV-1 Rev using a cell-based assay discovered PKF050-638, an *N*-azolylacrylate[396]. *N*-azolylacrylates were shown to potently inhibit XPO1-mediated nuclear export and to covalently bind the cysteine$_{528}$ residue of XPO1[396, 397]. Other small molecule XPO1 inhibitors have also been described and include S109, CBS9106, compound 3,4 and 5219668[389, 390]. As with all other XPO1 inhibitors described so far, these interact with XPO1 through covalent modification of cysteine$_{528}$.

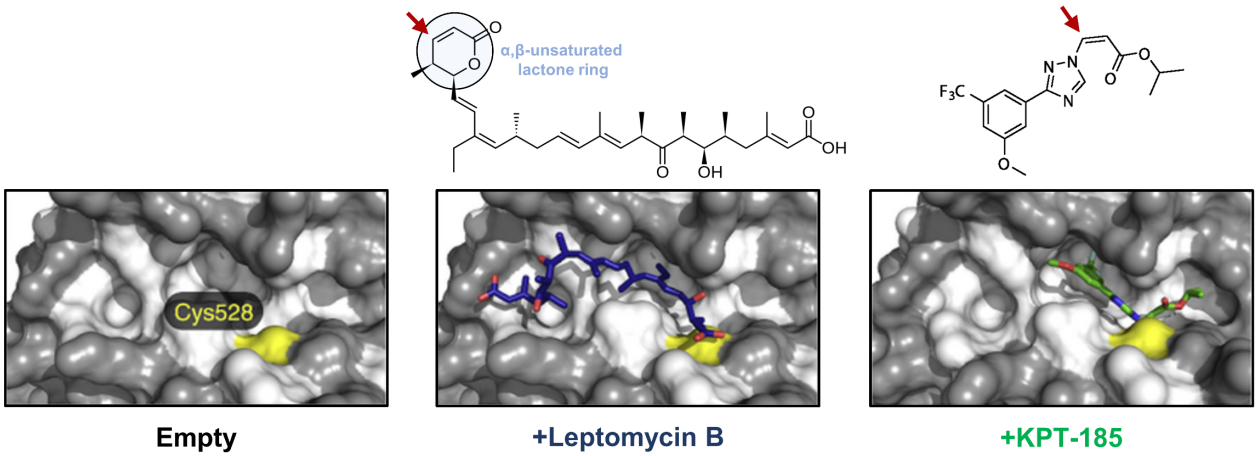


Figure 14: *binding of leptomycin B and the SINE KPT-185 to the NES-binding cleft (adapted from[374]).* Both leptomycin B and SINE interact covalently with the cysteine$_{528}$ residue located in the NES-binding cleft of XPO1. The chemical structure for leptomycin B and KPT-185 is shown and the site for covalent binding to cysteine$_{528}$ is denoted by a red arrow. When leptomycin B is bound to XPO1, its α,β-unsaturated lactone ring is hydrolyzed, linearizing the chemical structure and stabilizing the covalent bond with XPO1.

1.4.3.4 Improved small molecule inhibitors of nuclear export (SINE)

Based on the potential of using nuclear export inhibitors for the treatment of disease, Karyopharm Therapeutics set out to optimize the *N*-azolylacrylate scaffold of PKF050-638 *in silico*[398] using consensus induced fit docking based on the XPO1 crystal structure[332, 372]. These improved *N*-azolylacrylates (also called selective inhibitors of nuclear export, SINE) inhibit the formation of XPO1-cargo complexes. Co-crystal structures of the SINE with XPO1 demonstrated their binding to the cysteine$_{528}$ residue in the hydrophobic cleft of XPO1 (**Fig. 14** right panel)[397, 399]. Like leptomycin B, these improved inhibitors demonstrated potent activity against multiple types of cancer in various *in vitro* as well as *in vivo* models[359, 390, 399-407]. Although the mechanism of action of SINE is similar to leptomycin B, selinexor (KPT-330), the SINE clinical candidate, demonstrated decent tolerability with manageable side effects and high response rates as a single agent in phase 1 trials for human malignancies[408, 409]. Consequently, selinexor (KPT-330), has now entered phase 1-3 trials focused against solid and haematological malignancies (clinicaltrials.gov). These striking results might be explained by differences between leptomycin B and the SINE. First, SINE are small molecules that do not occupy the complete hydrophobic cargo-binding cleft of XPO1, whereas leptomycin B occupies almost the entire groove (**Fig. 14** compare the middle and right panels). This difference has been proposed to allow the accommodation of some cargos by XPO1 when bound by the SINE, although this hypothesis has yet to be clarified. Moreover, in contrast to leptomycin B[386], XPO1 binding by the SINE is not stabilized by hydrolysation, which has led to the suggestion that SINE binding of XPO1 is (slightly) more reversible than binding by leptomycin B[390]. The SINE also induce rapid proteasome-mediated degradation of XPO1 coupled to upregulation of *XPO1* mRNA[407], while leptomycin B does not or does so to at least a lesser extent. This suggests that SINE treatment leads to resynthesis of new XPO1, which further suggests that nuclear export inhibition by SINE is less permanent in the cell than treatment with leptomycin B.

CHAPTER II

Objectives of the research

Background

Most drugs exert their therapeutic effect by modulating the function of one or a few target proteins in a living cell or tissue. To achieve this modulation, drugs typically interact with their target through direct physico-chemical interactions with functionally important domains of the target protein. These interactions can lead to activation or inhibition of the target protein's function and the extent of engagement between the target protein and the drug dictates the drug's efficacy. Thus, understanding the target protein of a drug helps to predict possible side effects and the validity and tractability of the envisioned therapy. A detailed understanding of how a drug interacts with its target protein also supports hit-to-lead optimization and biomarker development (**Fig. 1**). Strong biomarkers related to the drug's target help predict patient response to drug treatment. As a result, it is important to uncover a drug's cellular target protein early in the drug discovery and development process.

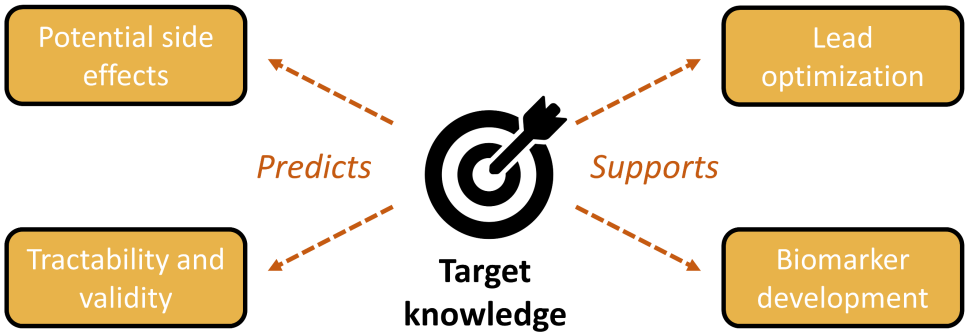


Figure 1: *understanding the molecular target of a drug supports the drug development process.*

Given the importance of understanding the mechanism of action of a candidate drug and its corresponding molecular target for the success of any drug discovery program, many target deconvolution and validation approaches have been developed over time. However, none of these methods is suitable for all circumstances and no universal solution exists. Therefore, uncovering the cellular target of small molecules is still one of the most challenging and time-consuming steps during the drug discovery process. The gold standard proof for a drug's target is the identification of mutations that confer resistance in a cellular context. For this reason, genetic approaches provide powerful tools for identification of a drug's cellular target. Unfortunately, applying such approaches in a mammalian context is still troublesome, although much progress has been made with the discovery and development of RNA interference and next-generation sequencing technologies. Current genetic methods are either plagued by technical issues, are not well suited to probe essential genes or require whole-exome sequencing combined with complex bio-informatics to deconvolute the relevant drug resistance conferring mutations. As a result, the field would greatly benefit from new approaches that can accelerate the drug resistance selection process and simplify subsequent identification of resistance mutations in a mammalian context. Now just a few years ago, a powerful technology for genetic manipulation had taken the stage: CRISPR-Cas-mediated genome editing. As the advent of the CRISPR technology coincided with the start of this PhD thesis, we wanted to explore the potential of this new and exciting technology. We set out to see whether CRISPR-Cas-mediated genome editing can be applied to validate and identify the cellular target protein of candidate anticancer agents still in development.

Aim 1

Our initial goal was to investigate the drug-target selectivity of clinical small molecule inhibitors of nuclear export. These molecules have entered clinical trials targeted against various solid and hematological malignancies and are believed to exert their anticancer activity through modulating the function of the Exportin-1 (XPO1/CRM1) protein. However, at the time it was unclear whether in human cells the anticancer activity of these inhibitors is selectively caused by inhibition of XPO1 function. As drug-target validation is an essential step in drug discovery, we aimed to validate this proposed drug-target selectivity using CRISPR/Cas9-induced homology directed repair of *XPO1* (**Fig. 2** right panel). The results of this validation study are described in **chapter III**.

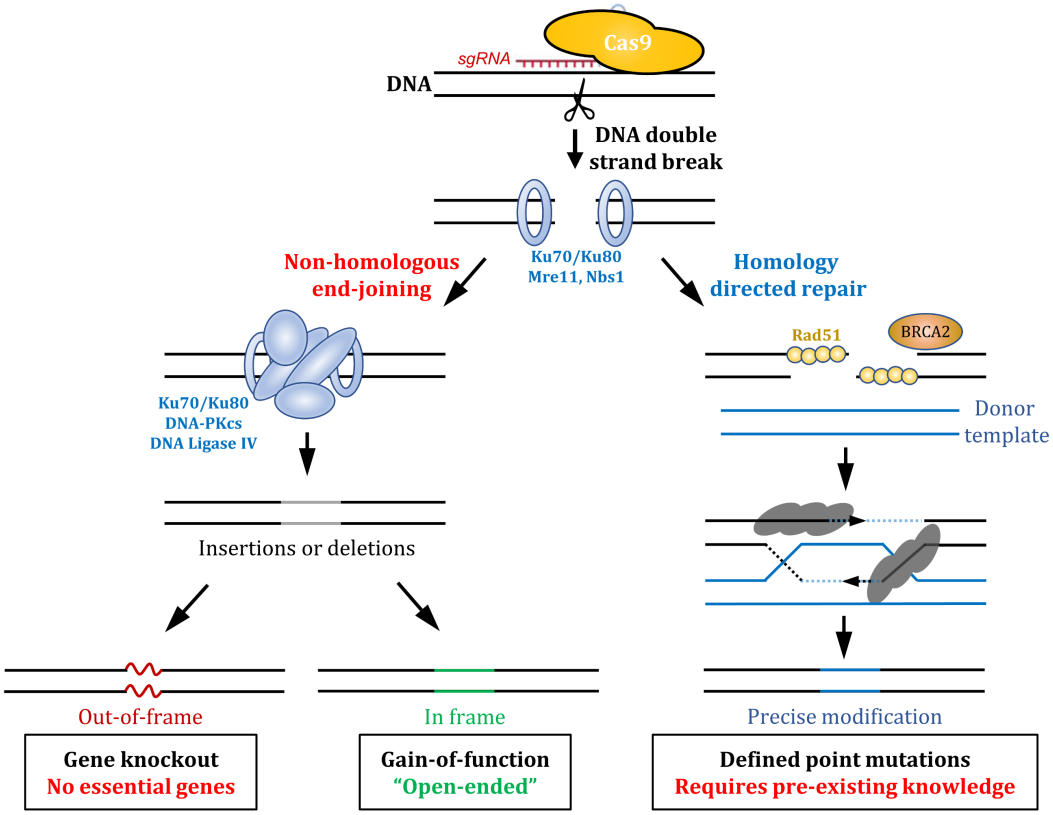


Figure 2: *overview of DNA repair outcomes of a CRISPR/Cas-induced DNA double strand break.*

Aim 2

For aim 2 we set out to develop a new target deconvolution method based on CRISPR/Cas-mediated genome editing. The gold standard proof for a drug's target is the identification of functional mutations that confer resistance in a cellular context. For this reason, genetic screens are powerful tools for drug mechanism of action studies. However, current screens for use in a human cellular context are either not well suited to identify essential genes or require whole exome sequencing combined with complex bio-informatics to deconvolute the relevant resistance mutations. Drawing a parallel to enhance the occurrence of drug resistance mutations by ultra-violet radiation, we hypothesized that targeted induction of NHEJ repair of DNA double strand breaks by CRISPR/Cas endonucleases can be utilized for protein mutagenesis to facilitate the formation of gain-of-function drug resistance mutations, even when mutagenizing essential genes (**Fig. 2** left panel). We set out to explore this principle using three anticancer agents for which the drug-target interactions are known. After validation of the concept, we established a new genetic target deconvolution approach utilizing large-scale CRISPR sgRNA tiling libraries to identify the cellular target of a novel anticancer agent still under development. The results of this study are described in **chapter IV**.

CHAPTER III

Article 1: validation of the drug-target selectivity of clinical XPO1 inhibitors

Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing

Jasper Edgar Neggers[1], Els Vanstreels[1], Thomas Vercruysse[1], Maarten Jacquemyn[1], Erkan Baloglu[2], Sharon Shacham[2], Marsha Crochiere[2], Yosef Landesman[2], Dirk Daelemans[1]

[1]KU Leuven Department of Microbiology and Immunology, Laboratory of Virology and Chemotherapy, Rega Institute for Medical Research, Herestraat 49, 3000 Leuven, Belgium.

[2]Karyopharm Therapeutics Inc, 85 Wells Ave, Newton, MA 02459, USA.

Abstract

Exportin-1 (CRM1/XPO1) is a crucial nuclear export protein that transports a wide variety of proteins from the nucleus to the cytoplasm. These cargo proteins include tumor suppressors and growth-regulatory factors and as a result XPO1 is considered a potential anticancer target. From this perspective, inhibition of the XPO1-mediated nuclear export by orally bioavailable selective inhibitors of nuclear export (SINE) has shown broad-spectrum anticancer activity. Furthermore, the clinical candidate SINE, selinexor (KPT-330), is currently in multiple phase 1/2/2b/3 trials for the treatment of cancer. However, it remains unclear if the anticancer activity of SINE compounds in a cellular context is selectively caused by inhibition of XPO1 function and not by other indirect mechanisms. Validation of drug-target interaction is essential in drug discovery and development and the ultimate proof for drug-target validation requires the introduction of mutations that confer resistance in a cellular context, an approach that is not straightforward in mammalian cells. Using CRISPR/Cas9 genome editing, we show a C528S substitution in endogenous *XPO1* confers cancer cells with resistance to selinexor. XPO1 mutant cells were resistant to the induction of cytotoxicity, apoptosis, cell cycle arrest, and inhibition of XPO1 function, including direct binding of the drug to XPO1. These results validate XPO1 as the cellular target of selinexor and highlight the selectivity of this drug toward the cysteine$_{528}$ residue of XPO1. Finally, our findings demonstrate that CRISPR/Cas9 genome editing enables drug-target validation and drug-target selectivity studies in cancer cells.

Adapted from the original publications:

Neggers et al. *Chemistry and Biology,* volume 22(1), pages 107-116, January 22 (2015)
doi: 10.1016/j.chembiol.2014.11.015

Neggers et al. *Oncotarget*, volume 7(42), pages 68842–68850, October 18 (2016)
doi: 10.18632/oncotarget.11995

Introduction

Human exportin-1 (XPO1), also known as chromosome region maintenance 1 protein (CRM1), is a key nuclear-cytoplasmic transport protein or karyopherin that exports a broad range of different cargo proteins out of the cell's nucleus[340-344, 410, 411]. These cargo proteins include tumour suppressor and growth regulatory related proteins, including NPM1, p53, IκB, pRb, BRCA1, APC and FOXO family proteins, various viral proteins and RNA binding proteins, and correct XPO1 function is key to normal cell homeostasis[331, 410, 411]. In recent years, overexpression or dysfunction of XPO1 has been observed in different types of cancer and has been correlated with poor prognosis, tumour size and resistance to therapy[357, 360, 361]. Indeed, alterations in XPO1 expression levels may cause subcellular mislocalization of tumour suppressor proteins and cell cycle regulators, resulting in uncontrolled cell growth and carcinogenesis[412-414]. In addition, a recurrent mutation in codon 571 of *XPO1* (p.E571K/G), located in the hydrophobic cargo-binding pocket has been reported in cancer, including chronic lymphocytic leukaemia, oesophageal squamous cell carcinoma and primary mediastinal large B cell lymphoma[364, 367, 368].

XPO1 has been considered a molecular target for cancer treatment for over 20 years and has resulted in several studies identifying small molecule inhibitors of XPO1. Up until a few years ago, Leptomycin B (LMB) was the only XPO1 inhibitor that was investigated in the clinic. However, LMB was discontinued due to toxic side effects including nausea, gastrointestinal effects and fatigue[394]. More recently, *N*-azolylacrylate small-molecule inhibitors of the XPO1-mediated nuclear export[396, 397] were rationally optimized *in silico*[398] based on the XPO1 crystal structure[332, 372]. These improved *N*-azolylacrylate small-molecule inhibitors of nuclear export, called SINE, demonstrate potent activity against multiple types of cancer and have been shown to induce apoptosis and to abolish cancer growth in various *in vitro* as well as *in vivo* models of cancer[359, 399-407]. Importantly, selinexor (KPT-330), the clinical candidate of the SINE molecules currently being tested in various clinical studies, demonstrated high response rates as a single agent in phase 1 trials for heavily pre-treated, relapsed and refractory haematological and solid tumour malignancies in humans[408, 409] and is currently being evaluated in combination with other chemotherapeutic agents. Preclinical analogues of this drug inhibit the formation of XPO1-cargo complex *in vitro* and co-crystal structures with XPO1 protein have demonstrated their binding into the cargo-binding pocket of XPO1[399, 404]. The binding is proposed to involve an irreversible Michael addition type interaction of their acrylate moiety with the cysteine$_{528}$ residue of XPO1[397, 399]. These SINE compounds have been shown to cause an accumulation of tumour suppressor proteins in the nucleus of treated cells, which is correlated with the induction of cell cycle arrest and apoptosis[359, 399-407]. However, the specificity of this drug-target interaction in mammalian cells has not been validated. Indeed, it has not been demonstrated that the anticancer activity of SINE in a cellular context is selectively caused by inhibition of XPO1 function and not by other mechanisms resulting in or adding to the observed anticancer activity. To select for resistance and to better understand the mechanism of action of SINE compounds, HT-1080 cancer cells have been treated *in vitro* with gradually increasing concentrations of selinexor[415]. However, resistance was reached only after 10 months of continuous exposure and resistant cells displayed an altered expression pattern in several key signaling pathways, including upregulation of anti-apoptotic and downregulation of pro-apoptotic genes, while no mutations were detected in XPO1.

Drug-target validation is an essential step in drug discovery. While identification of drug-resistance mutations is regarded as the 'gold-standard' for target confirmation, further validation of drug-target interaction requires the resistance mutation to be introduced into the wild-type background, a methodology widely used for prokaryotes, lower eukaryotes and viruses. However, until the advent of the CRISPR/Cas9 genome editing technique[223-225], this approach was not straightforward in mammalian cells. Very recently, proof-of-concept for drug validation in human cells has been shown based on resistance selection and genome sequencing

in combination with CRISPR/Cas9 genome editing[120, 121]. We reasoned that if SINE are selective for XPO1, genomic mutation of the cysteine$_{528}$ residue in XPO1 should cause resistance against these compounds. However, in higher eukaryotes, this residue is very well conserved[353, 373], raising the question whether a genomic mutation of this cysteine$_{528}$ residue is possible in human cells. Interestingly, variations of this cysteine residue are rarely observed in lower eukaryotes. For example, *Saccharomyces cerevisiae* and *Aspergillus nidulans* both contain a threonine residue at this position, while a mutant strain of *Shizosaccharomyces pombe* has been reported to carry a serine residue[387].

In this study we used CRISPR/Cas9 genome editing in combination with homology directed repair (HDR) to introduce an XPO1 C528S mutation in human cancer cell lines. These mutant cell lines were then used to validate XPO1 as the specific cellular target for the SINE compounds, including the clinical stage XPO1 inhibitor selinexor, and to assess the selectivity of this drug for the cysteine$_{528}$ residue in the hydrophobic cargo-binding groove of XPO1.

Results

Generation of XPO1^{C528S} knock-in mutants by CRISPR/Cas9 genome editing

To obtain cell lines carrying a cysteine to serine substitution at position 528 inside the XPO1 hydrophobic cargo-binding pocket, we employed CRISPR/Cas9 genome editing to alter the corresponding TGT DNA codon to TCA using homologous recombination. HL-60, Jurkat and K-562 leukaemia cells, all sensitive to treatment with SINEs, were co-transfected with a plasmid expressing SpCas9 fused to a nuclear localization signal sequence, a plasmid expressing a CRISPR guide RNA (sgRNA) targeted against *XPO1* and a 135 bases long single-stranded oligodeoxynucleotide (ssODN) repair donor template containing the TCA nucleotide mutation and three silent mutations to introduce a serine residue at position 528 in the hydrophobic cargo-binding pocket of XPO1 (**Fig. 1a**). We hypothesized if the C528S mutation in XPO1 *per se* is not lethal, XPO1^{C528S} mutants may be insensitive to KPT-185, the preclinical analogue of selinexor (KPT-330) (**Fig. 1b**). As a result, to enrich for cells that underwent homology-directed repair, transfected cells were treated with 100 nM of KPT-185 for 3 consecutive days. Only few cells survived (<10%), indicating a low efficiency of homology-directed repair. The number of surviving cells was dependent on transfection efficiency and cell-type, as more cells survived in the easy to transfect K-562 and HL-60 cell lines (data not shown). Following transfection, single cell derived colonies were obtained. From these colonies genomic DNA was extracted and exon 15 of the *XPO1* gene was amplified by PCR and then sequenced using Sanger sequencing. Most clones integrated the desired missense mutation at only one of the *XPO1* alleles, while other alleles contained either the wild-type sequence, the silent mutations only, but not the desired missense mutation, or an insertion or deletion caused by non-homologous end-joining (NHEJ) (classified as hemizygous) (**Fig. 1c**, **Table 1, Supplementary Table 1**). The remaining clones integrated the TGT to TCA mutation in both alleles. Two K-562 colonies only contained the wild-type sequence, suggesting they tolerated the initial low dose selection (100 nM) with KPT-185. All sequences containing the desired missense TCA mutation (C528S) also contained the three additional silent mutations, effectively ruling out spontaneous generation of resistance mutations during drug selective pressure. One of the two homozygous clones for the Jurkat cell line (clone 6) was selected for further validation. This clone showed a slightly higher XPO1 expression level compared with parental wild-type cells when analysed by immunoblot (**Fig. 1d**), quantitative RT-PCR (**Fig. 1e**), and confocal immunofluorescence microscopy (**Fig. 1f**).

Cell line	Leukemia	Substitution	Homozygous	Heterozygous	Hemizygous	Wild-type	Clones
HL-60	**APML**	TGT$_{528}$->TCA	21 (21.9%)	22 (22.9%)	53 (55.2%)	0 (0%)	**96**
Jurkat	**T-ALL**	TGT$_{528}$->TCA	2 (12.5%)	10 (62.5%)	4 (25%)	0 (0%)	**16**
K-562	**CML**	TGT$_{528}$->TCA	24 (33.3%)	11 (15.3%)	35 (48.6%)	2 (2.8%)	**72**

Table 1: *overview of the obtained XPO1 mutant clones after CRISPR/Cas9 genome editing.*

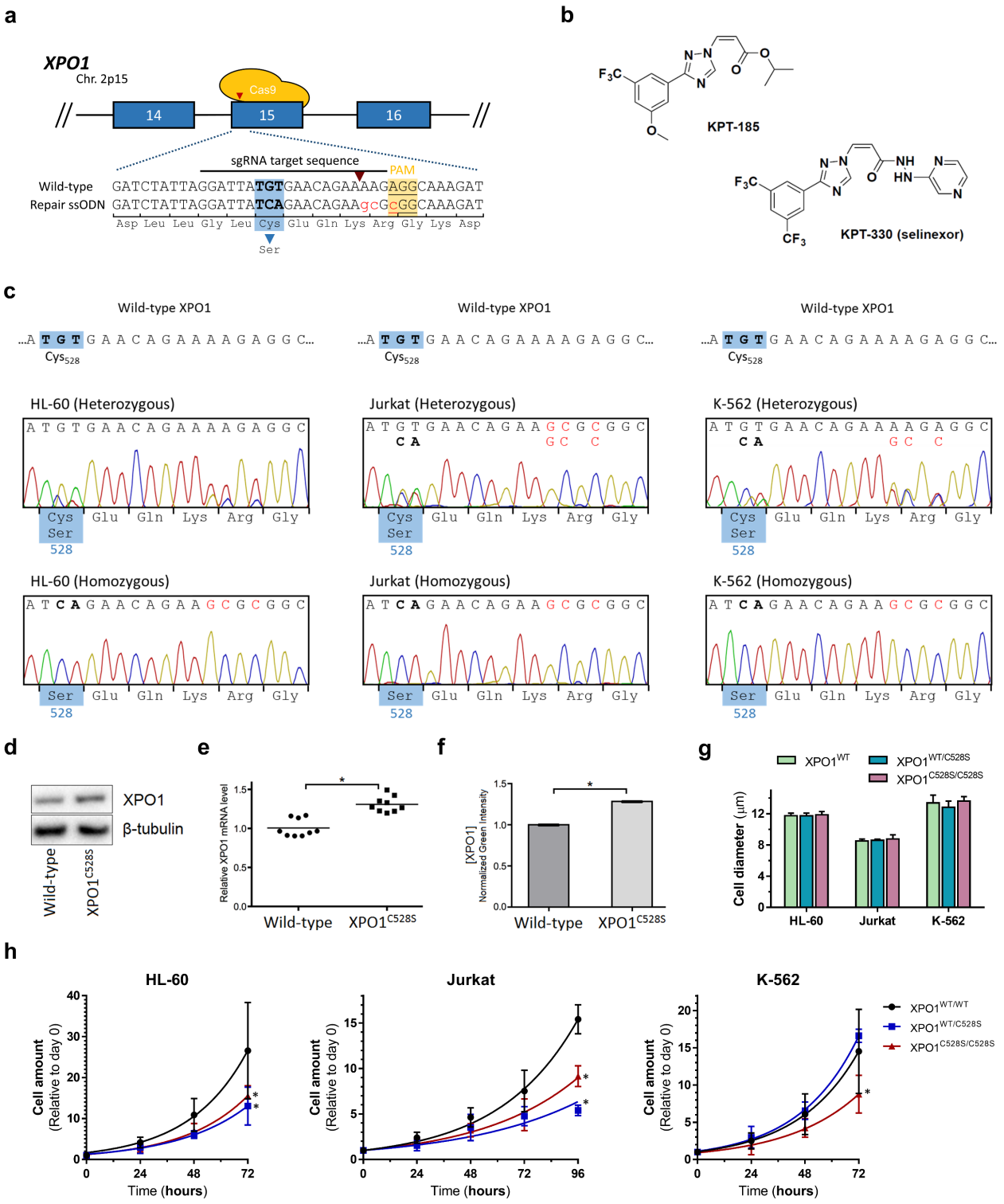


Figure 1: *generation of XPO1^{C528S} cell lines by CRISPR-Cas9 induced homologous recombination.*
a. Schematic overview of CRISPR/Cas9-induced homologous recombination of human *XPO1*. The nucleotide sequence for both the wild-type *XPO1* and the repair donor template are shown with the cysteine$_{528}$ codon and NGG PAM site highlighted. The sgRNA target site is shown, with the SpCas9 cut site denoted by a red arrowhead. Silent mutations are shown in red. ssODN: single-stranded oligodeoxynucleotide; PAM: protospacer adjacent motif.
b. The chemical structures of the SINEs KPT-185 and KPT-330 (selinexor).
c. Sequencing chromatograms of *XPO1* in hetero- and homozygous XPO1^{C528S} clones.
d. Visualization of XPO1 protein expression in wild-type and homozygous XPO1^{C528S} (clone 6) Jurkat cells by Western-blot. β-tubulin was used as internal loading control.

e. Relative comparison of mean *XPO1* mRNA expression levels in wild-type and homozygous XPO1^{C528S} (clone 6) Jurkat cells by quantitative PCR (unpaired student's t-test p-value: <0.0001); *GAPDH* and *β-actin* were used as internal controls.
f. Relative comparison of mean XPO1 protein expression in wild-type (2,269 cells) and homozygous XPO1^{C528S} (1,866 cells, clone 6) Jurkat cells as measured by immunostaining and quantified by confocal immunofluorescence microscopy (unpaired student's t-test p-value: <0.0001). Error bars indicate the 95% confidence interval.
g. Quantification of the cell diameter of wild-type and heterozygous or homozygous XPO1^{C528S} cell lines. Cell diameters were obtained with a LUNA II automated cell counter. Bars represent means and error bars indicate standard deviation obtained from 3 experiments. Averages were statistically compared by ANOVA, but no significant differences were determined.
h. Quantification of cellular growth rates from parental wild-type or heterozygous and homozygous XPO1^{C528S} mutant cell lines as determined with the CyQuant Direct Cell Proliferation kit. Cellular growth was analysed over a period of 3-4 days and normalized to day 0. Data points represent means and error bars indicate the standard deviation obtained from 2 (HL-60 and K-562) or 3 (Jurkat) experiments performed in triplicate. Exponential growth curves were obtained by fitting and statistical differences between the wild-type and mutant cell lines were determined with GraphPad Prism (ANOVA *p-value < 0.0001).

Interestingly, although heterozygous and homozygous XPO1^{C528S} mutant clones were viable and could be cultured for several months, mutant cells appeared to grow slower. To investigate this phenotype, the cell size of hetero- and homozygous mutants and parental cells were compared, but no differences were detected (**Fig. 1g**, p-values between 0.4 and 0.9). However, when the cell proliferation of wild-type, heterozygous and homozygous mutants was analysed over a period of 3-4 days (**Fig. 1h**) a slower growth rate, especially in homozygous mutants, was observed for mutant Jurkat and HL-60 cell lines (p<0.0001). Heterozygous K-562 mutant cells did not grow significantly slower (p=0.0968), although homozygous K-562 mutant cells clearly did (p<0.0001).

Serine substitution of C528 in XPO1 confers resistance to SINE compounds

The mutant XPO1^{C528S} cell lines that we have generated provide a powerful tool to directly investigate and validate XPO1 as the key cellular target for the clinical SINE compound KPT-330 (selinexor). Therefore, the cytotoxic effect of KPT-330 was assessed on parental wild-type and mutant XPO1^{C528S} cells. In addition, to investigate whether resistance to KPT-330 is recessive or dominant for the cysteine$_{528}$ mutation, both hetero- and homozygous mutants were treated. A remarkable reduction in cell viability of wild-type K-562, HL-60 and Jurkat cells was observed at nanomolar concentrations, while all the heterozygous and homozygous mutant cell lines were resistant to selinexor up to micromolar concentrations (**Table 2; Fig. 2a**). As heterozygous mutation was sufficient for drug resistance, the C528S drug resistance mutation is dominant. Interestingly, there was a slight difference in drug sensitivity between Jurkat hetero- and homozygous mutant cell lines that was not observed in the other cell lines.

	EC_{50} (nM)		
Cell Line	XPO1$^{WT/WT}$	XPO1$^{WT/C528S}$	XPO1$^{C528S/C528S}$
HL-60	126 ± 23	12,7364 ± 6,360	12,362 ±6,360
Jurkat	53 ± 4	6,490 ± 2,582	16,987 ± 1,052
K-562	184 ± 26	55,360 ± 13,836	38,575 ± 5,048

Table 2: *EC_{50} values for KPT-330 (selinexor).* Values were obtained from the viability assays (**Fig. 2a**) and represent means ± s.d. determined from 3 experiments performed in triplicate.

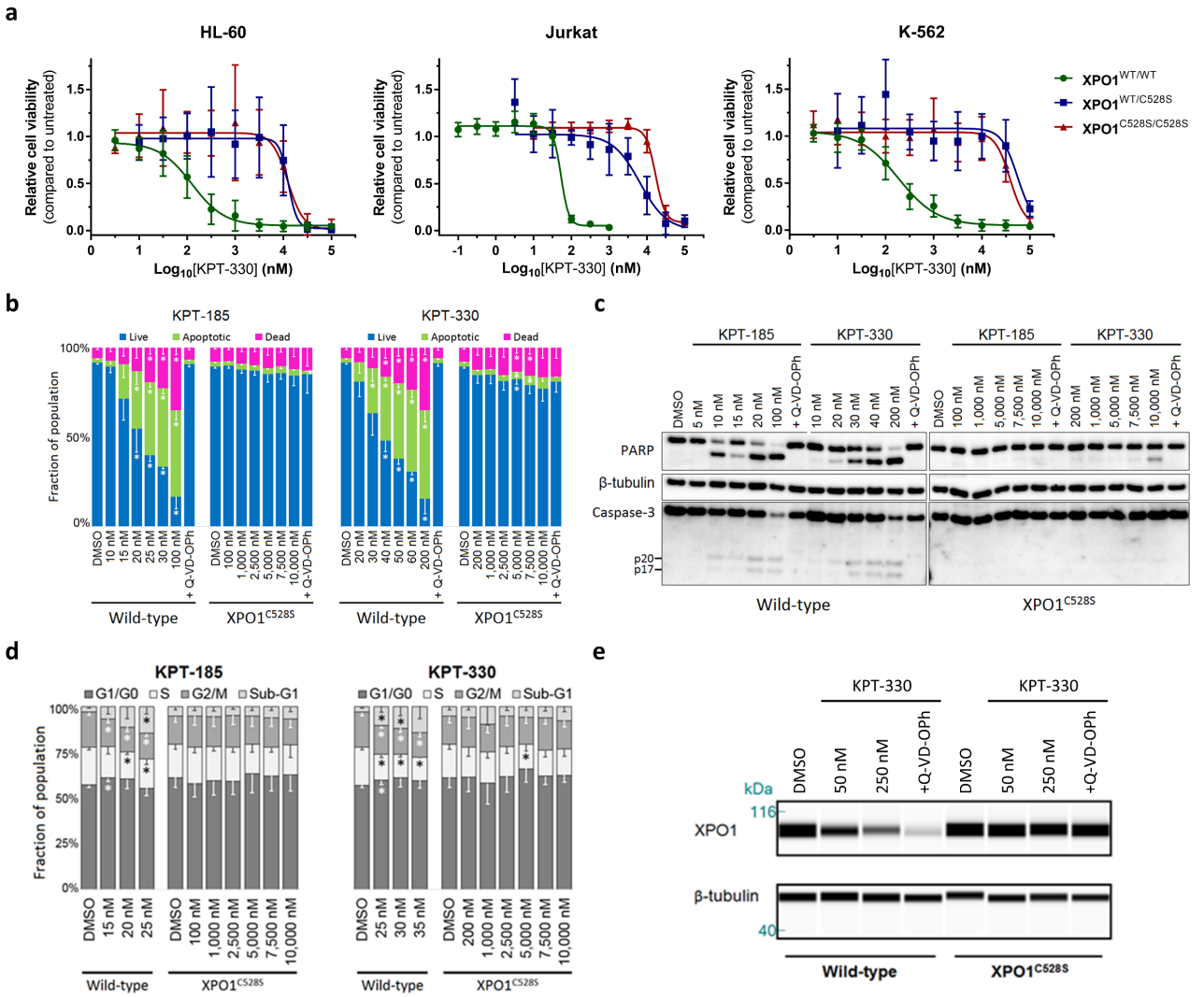


Figure 2: *serine substitution of cysteine$_{528}$ in XPO1 confers drug resistance to SINE.*
a. Cell viability of the different cell lines after 72hr of treatment with different concentrations of KPT-330 (selinexor). Points represent the average and error bars indicate standard deviation obtained from three independent experiments performed in triplicate.
b. Overview of apoptosis and cell death profiles in wild-type (parental) and homozygous mutant XPO1^{C528S} (clone 6) Jurkat cells as determined by flow cytometry with Annexin V and propidium iodide (PI) staining after 12hr treatment with carrier (DMSO) or compound (KPT-330 and KPT-185). Stacked columns represent 100%. Where indicated, the pan-caspase apoptosis inhibitor Q-VD-OPh was added in addition to the highest concentration of compound. Values represent averages ± s.d. determined over 3 experiments. The relative amount of apoptotic, dead or living cells in each condition was compared with the respective fraction of the DMSO control (unpaired student's t test ∗p value, <0.05).
c. Detection of caspase-3 and PARP cleavage by Western-blotting in wild-type and homozygous mutant XPO1^{C528S} (clone 6) Jurkat cells after 16hr treatment with KPT-185, KPT-330 or DMSO. Where indicated, the pan-caspase inhibitor Q-VD-OPh was added to the highest concentration of compound. β-tubulin served as internal loading control.
d. Overview of the cell cycle profiles as determined by PI staining in wild-type and XPO1^{C528S} (clone 6) Jurkat cells after 24hr treatment with carrier (DMSO) or various concentrations of KPT-185 or KPT-330. Stacked columns represent 100%. Values represent averages ± s.d. determined over 5 experiments. The fraction of each cell cycle phase in each condition was compared with the DMSO-treated control condition (unpaired student's t test, $*p < 0.05$).
e. Detection of KPT-330-induced degradation of XPO1 by Simple Western after overnight treatment of wild-type or homozygous C528S mutant (clone 6) Jurkat cells. Q-VD-OPh was added together with 250 nM KPT-330. β-tubulin served as internal loading control.

As the SINE compounds have been shown to rapidly promote apoptosis in cancer cell lines *in vitro*[359, 399-407] as well as in human trials[408], we wanted to investigate the apoptotic effects of KPT-185 and KPT-330 treatment of C528S mutant cells. For this purpose, the homozygous Jurkat XPO1^{C528S} mutant clone 6 was treated for 12 hours with both compounds and analysed using annexin V/PI staining coupled to flow cytometry. The homozygous clone was resistant up to micromolar concentrations of both compounds, while only low nanomolar concentrations of compound were sufficient to induce apoptosis in wild-type Jurkat cells (**Fig. 2b**). Only at extreme high concentrations of both compounds (>5 μM) the fraction of dead and apoptotic XPO1^{C528S} mutant cells appeared to increase slightly (**Fig. 2b**). The induction of apoptosis and cell death in wild-type cells by both compounds was caspase dependent as it was prevented by the addition of the pan-caspase apoptosis inhibitor Q-VD-OPh (**Fig. 2b**). Apoptosis is typically marked by activation of the caspase-3 proenzyme through proteolytic cleavage and inactivation of Poly (ADP-ribose) polymerase (PARP) through proteolytic activity by caspase-3 and 7. Western-blot analysis showed treatment of wild-type cells induced cleavage of caspase-3 and PARP at low nanomolar concentrations, while the C528S substitution conferred cells with resistance to caspase-3 activation up to micromolar concentrations (**Fig. 2c**).

Besides induction of apoptosis, treatment of cells with XPO1 inhibitors is known to arrest cell-cycle progression[359, 400, 402-405, 407]. To further validate that the cytotoxic effects induced by the SINE compounds are caused by the inhibition of XPO1, we next assessed the impact of the C528S mutation on the effect of KPT-185 and KPT-330 on cell cycle progression (**Fig. 2d**). Treatment of wild-type cells resulted in a slight, but significant, G1/G0 cell cycle arrest with lowered S and G2 phases after 24 hours of incubation with low concentrations of compound (**Fig. 2d**). An increase in sub-G1 apoptotic wild-type cells was also observed and becomes even more pronounced at higher compound concentrations, consistent with the levels of Annexin V staining (**Fig. 2b**). In contrast, treated homozygous Jurkat mutant XPO1^{C528S} cells showed almost no significant differences in the cell cycle profile as compared to the negative control condition at the same or even micromolar concentrations (**Fig. 2d**). Another phenomenon known to occur in SINE-treated cells is proteasome-mediated degradation of XPO1[407]. To examine whether the C528S mutation confers resistance against this induced XPO1 degradation, XPO1 protein was detected by Western-blotting of wild-type and homozygous XPO1^{C528S} mutant (clone 6) Jurkat cells after overnight treatment with KPT-330. As expected, XPO1 protein levels were decreased in wild-type cells after treatment, even in Q-VD-OPh treated cells, while XPO1^{C528S} mutant cells were protected against KPT-330-induced degradation of XPO1, further highlighting the C528S mutation confers resistance against treatment with SINE compounds (**Fig. 2e**).

The C528S mutation rescues XPO1 function from SINE-induced inhibition

To further confirm that the C528S mutation rescues XPO1 from inhibition by the SINE compounds, we assessed the effect of the compounds on nuclear export function of hetero- and homozygous mutant XPO1^{C528S} Jurkat cells. For this purpose, the subcellular localization of the endogenous RanBP1 XPO1 cargo protein was visualized in the presence of the compounds by immunostaining coupled to confocal fluorescence microscopy. In the absence of XPO1 inhibitor, RanBP1 was found in the cytoplasm of both wild-type and mutant cells (**Fig. 3a** top panels). Within three hours after the addition of compound, RanBP1 accumulated in the nucleus of wild-type cells as caused by inhibition of XPO1-mediated nuclear export (**Fig. 3a** left panels). Interestingly, the SINE compounds were unable to inhibit the nuclear export of RanBP1 in homozygous XPO1^{C528S} cells after 3 hours of treatment (**Fig. 3a** right panels), suggesting mutated XPO1^{C528S} is active and can bind cargo proteins. Indeed, it has previously been shown *in vitro* that mutant XPO1^{C528S} interacts with cargo proteins[373]. In addition, a cysteine to serine substitution has been found in mutants of the yeast *Schizosaccharamyces pombe*[387], which is dependent on functional XPO1 for survival. To quantify nuclear export of RanBP1 in the

presence of different concentrations of inhibitor, a minimal number of 200 cells were analysed for the ratio of nuclear over cytoplasmic RanBP1 staining of wild-type and XPO1 mutant Jurkat cells (**Fig. 3b**). Remarkably, the heterozygous mutant showed a similar resistance profile as the homozygous mutant. These findings further strengthen the conclusion that heterozygous mutation of the cysteine$_{528}$ residue is sufficient for drug resistance against the SINE. To further confirm mutant XPO1's ability to interact with cargo proteins in human cells and overcome inhibition by SINE compounds, we assessed XPO1-cargo binding in cells by means of a co-transfection assay with fluorescently tagged XPO1 variants and the prototype XPO1 cargo protein HIV-1 Rev[341]. HeLa cancer cells were co-transfected with plasmids expressing Rev fused to the blue fluorescent protein (BFP) and either wild-type XPO1 or mutant XPO1^{C528S} fused to the yellow fluorescent protein (YFP) (**Fig. 3c**)[416]. The Rev-BFP protein localized to the nucleoli of the cells while XPO1 was mainly found at the nuclear membrane and throughout the nucleus but was generally excluded from the nucleoli (**Fig. 3c**, panels a and b). When both proteins are co-expressed, a significant portion of wild type or mutant XPO1-YFP protein co-localized with Rev-BFP in the nucleoli (**Fig. 3c**, panel c) [417]. It must be noted that co-localization of two proteins inside cells is generally not a direct proof of interaction. However, as Rev is known as the prototype cargo for XPO1, and because the XPO1 protein is excluded from the nucleoli in the absence of Rev and only localized to the nucleoli when Rev was present, we believe the observed co-localization is the result of an interaction between XPO1 and Rev. An additional argument is that the co-localization with wild-type XPO1 was disrupted upon treatment with KPT-330 (**Fig. 3c** panel d), which is in agreement with biochemical assays where preclinical analogues of KPT-330 have been shown to disrupt the interaction between wild-type XPO1 and the Rev nuclear export signal sequence[399, 404]. Moreover, upon treatment with KPT-330, the interaction between Rev-BFP and XPO1^{C528S}-YFP was not affected by KPT-330, demonstrating the C528S substitution confers resistance to the drug with respect to cargo binding.

Next, the interaction of SINE compounds with wild-type XPO1 or mutant XPO1^{C528S} cellular protein was assessed. For this purpose, wild-type and hetero- or homozygous XPO1^{C528S} mutant Jurkat cells were treated with a biotinylated analogue of KPT-276 (KPT-9058) (**Fig. 3d**) to pull down XPO1 protein. KPT-276 is another SINE compound and a structural analogue of KPT-330 and KPT-185 carrying the same acrylate warhead. It shows similar anticancer activity as KPT-185 and KPT-330[400, 403, 407]. Like KPT-185 and KPT-330, this biotinylated derivative induced potent cytotoxic effects in wild-type Jurkat cells (EC_{50} of 187.3 ± 9.8 nM), while it was inactive against the homozygous XPO1^{C528S} mutant cells (EC_{50} of 18.5 ± 2.3 μM) (**Fig. 3e**). Streptavidin affinity chromatography was used to isolate the compound from the cell lysate, and coprecipitated XPO1 protein was detected by Simple Western analysis (**Fig. 3f**). As expected, XPO1 protein was pulled down from wild-type cells, while it was impossible to extract XPO1 protein from homozygous XPO1^{C528S} mutant cells. Quantification of the pulled down protein from heterozygous XPO1^{C528S} mutants showed the amount of pulled down XPO1 protein was about half of that pulled down from wild-type cells (**Fig. 3g**), confirming half of the expressed XPO1 in the heterozygous mutants consists of wild-type protein while the other half consists of mutant XPO1^{C528S} protein. Altogether these results validate XPO1 as the target for the SINE compounds and more specifically pinpoint the cysteine$_{528}$ residue as a highly selective anchor point for these drugs.

Finally, a recurrent mutation at codon 571 (p.E571K/G) in the *XPO1* gene has recently been described in some leukemia patients[367, 418]. Because this mutation is located inside the hydrophobic cargo-binding pocket near cysteine$_{528}$ of XPO1, we evaluated whether this mutation could affect KPT-330 activity. Therefore, we edited the *XPO1* gene of HAP1 chronic myelogenous leukemia cells using CRISPR/Cas9 to introduce the E571K mutation. Sensitivity of these mutant cells to KPT-330 treatment was measured and compared to wild-type cells (**Supplementary Figure 1**). Interestingly, wild-type and mutant cells were equally sensitive to KPT-330, highlighting this mutation does not affect binding of the drug to XPO1.

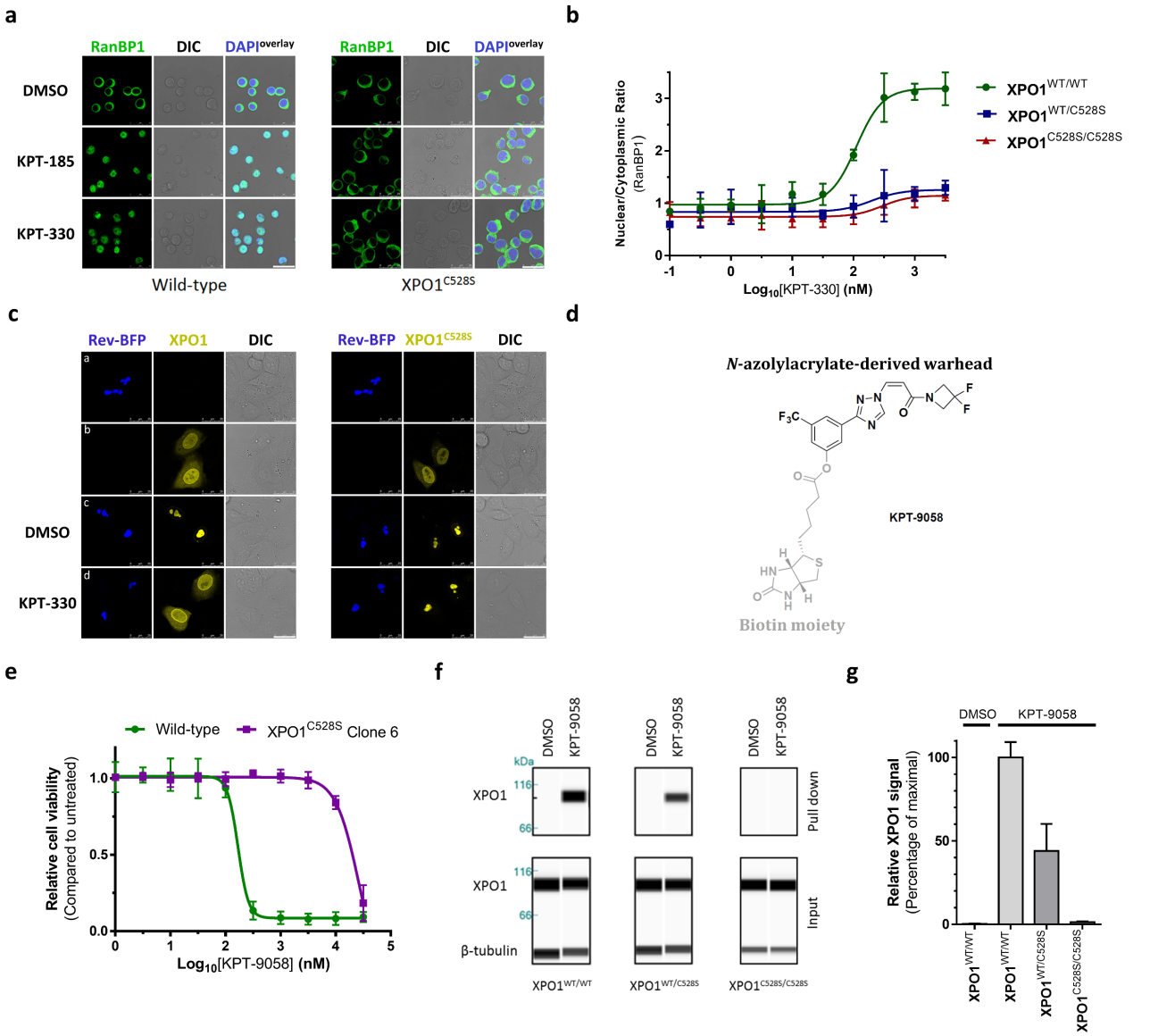


Figure 3: *XPO1^{C528S} substitution protects XPO1 function from SINE and prevents binding of SINE to XPO1.*
a. Immunofluorescence analysis of RanBP1 cargo (green) in wild-type and mutant XPO1^{C528S} (clone 6) Jurkat cells after 3hr treatment with DMSO or 1 μM compound (KPT-185 or KPT-330). Cell nuclei are counterstained with DAPI (blue). DIC: differential interference contrast.
b. Quantification of the XPO1-mediated nuclear export in Jurkat cells showing the nuclear/cytoplasmic localization of the cargo protein RanBP1 in the presence of increasing concentrations of KPT-330. RanBP1 was visualized (green) with a high content imaging system by immunostaining. Cell nuclei were counterstained with DAPI (blue). Values represent means and error bars indicate standard deviation obtained from 2 experiments. At least 200 cells were counted per experiment for each data point.
c. Confocal fluorescence microscopy imaging of colocalization of wild-type XPO1-YFP or mutant XPO1^{C528S}-YFP (yellow) with Rev-BFP (blue) in HeLa cells. Colocalization of XPO1 with Rev was visualized one hour after the addition of DMSO or 1 μM KPT-330.
d. Chemical structure of KPT-9058, a biotinylated analogue of KPT-276.
e. Cell viability of wild-type and mutant XPO1^{C528S} (clone 6) Jurkat cells after treatment for 72 hours with increasing concentrations of KPT-9058. Data points show averages ± standard deviation collected over 3 independent experiments performed in singular.
f. Pull down of XPO1 protein with KPT-9058. Jurkat cells were incubated with DMSO or 1 μM of KPT-9058 for 2hr before streptavidin affinity chromatography was used to pull down XPO1 protein.
g. Quantification of XPO1 protein levels in streptavidin-purified extractions using KPT-9058 as shown in panel f. Values were divided by the signal of β-tubulin and compared relatively to the corrected XPO1 signal in wild-type cells. Bars represent means and error bars indicate standard deviation. The experiment was performed twice.

Discussion

Drug-target validation is an indispensable step in drug discovery. While identification of drug-resistance mutations is regarded as the ultimate proof for target confirmation, further validation of drug-target interaction requires the resistance mutations to be introduced into the wild-type background. This genetic methodology is widely used in prokaryotes, lower eukaryotes and viruses, but has not been straightforward in mammalian cells. Selinexor (KPT-330) is the clinical candidate of a novel promising class of anticancer drugs targeting XPO1 (SINE) and has displayed promising preliminary results *in vitro* and in human clinical trials[408, 409, 419]. The drug has been shown to interact with the cysteine$_{528}$ residue located in the hydrophobic cargo-binding pocket of XPO1[420]. However, although previous *in vitro* studies have demonstrated the interaction of these molecules with XPO1 and the accumulation of tumour suppressor proteins in the nucleus of treated cells is correlated with the induction of apoptosis, the selectivity of selinexor for XPO1 and the direct causality between inhibition of XPO1 and the observed anticancer mechanism of action have not been demonstrated in human cells. Here, we applied CRISPR/Cas9 genome editing to validate the anticancer activity of selinexor is selectively caused by its interaction with XPO1 in human cancer cells.

Using CRISPR/Cas9 genome editing combined with homology-directed repair, we obtained hetero- and homozygous leukaemia cell-lines containing a C528S substitution in XPO1. We then evaluated the effect of selinexor (KPT-330) on cell viability, apoptosis and cell cycle progression in these mutant cells as these parameters were found greatly affected by this drug in both preclinical models of cancer as well as in human trials. Our mutant XPO1^{C528S} cells were highly resistant to selinexor when compared to wild-type with the C528S substitution conferring >250-fold resistance to selinexor. Interestingly, heterozygous mutation of cysteine$_{528}$ was sufficient to elicit complete resistance to selinexor, demonstrating the dominant nature of this resistance mutation. Furthermore, this dominant resistance also demonstrates 50% of the total XPO1 protein pool is sufficient to maintain cellular homeostasis within the cell when treated with XPO1 inhibitors. However, substitution of the cysteine residue by serine appears to affect the cell proliferation slightly, which may be caused by an altered affinity of mutant XPO1^{C528S} to a subset of cargo-proteins. This might alter the nuclear-cytoplasmic distribution of these XPO1 cargo proteins, which in turn could disrupt their function. We also observed a slightly increased XPO1 expression in one of the homozygous mutant cell-lines, although we believe this increase does not significantly contribute to the observed >250-fold resistance. Besides selinexor's reduced effect on cell viability, selinexor did not induce cell cycle arrest nor inhibit XPO1-mediated cargo export in mutant cells. However, at very high concentrations, a small increase in mutant apoptotic cells could be observed. By using a biotinylated analogue, we demonstrated the direct binding of this class of compounds to cellular XPO1 protein while the single C528S substitution conferred cells with resistance to this drug-target interaction. Together, these results demonstrate that selinexor's anticancer activity is caused by selective inhibition of XPO1 function and highlight the unique selectivity of this class of XPO1 inhibitors for the cysteine$_{528}$ residue of XPO1 in human cells.

Little is known on possible resistance mechanisms to selinexor *in vivo*. Several studies focused on patient samples have suggested that recurrent mutation of residue E571 located in the NES binding groove of XPO1 might have prognostic or therapeutic value for different types of cancer[358, 365-367, 421]. Interestingly, we showed the E571K mutation does not alter sensitivity of isogenic HAP1 cells to selinexor. In addition, this mutation also does not alter drug sensitivity in primary mediastinal B-cell lymphoma[365]. A cryptic fusion of *XPO1* with *MLLT10* has been reported in T-ALL[422], but the effect of this translocation on selinexor activity remains to be investigated. Additionally, selinexor is currently under investigation in many different clinical trials focused on its anticancer activity, but to our knowledge none of these trials have examined the appearance of cysteine$_{528}$ mutations in the different treatment regimens. Although this cysteine residue is well conserved among higher eukaryotes[373, 387], we were able

to introduce the C528S point mutation into human leukaemia cells. This demonstrates that the cysteine residue inside the hydrophobic cargo-binding groove of XPO1 can be effectively mutated in higher eukaryotes. Intriguingly, this mutation also resulted in decreased cellular growth rates, showing this mutation might reduce cancer cell fitness. We have chosen to introduce a serine residue because of its chemical similarity to cysteine and because this mutation has been found in some yeast strains[387], but other amino acid substitutions can be envisaged. However, it remains to be seen whether a heterozygous or homozygous mutation of the cysteine residue can arise in the clinic and further *in vivo* studies are needed to elucidate the possible effect of C528 substitutions in XPO1 on cellular homeostasis, cancer cell fitness, tumor growth rates and selinexor's anticancer activity.

Our findings illustrate the general utility of introducing point mutations in the genome of mammalian cells to validate drug-target interactions and investigate target selectivity. Two studies recently provided proof-of-concept for the use of CRISPR/Cas9-mediated genome editing for drug-target validation[120, 121]. Both studies sequenced the transcriptome of drug-resistant clones to guide their validation experiments. Our study shows the CRISPR/Cas9 HDR approach allows for the determination of target selectivity of drugs for which drug-resistant clones are non-existent and therefore already provides important predictive information on the resistance profile that could arise against a drug. Furthermore, we managed to obtain homozygous mutants demonstrating proof-of-principle for CRISPR/Cas9 genome editing to study recessive drug-resistance mutations.

Although the CRISPR/Cas9 technology is sometimes criticized for its off-target genome alterations[216, 217, 244], we did not screen for possible off-target mutations induced by the CRISPR/Cas9 technology. We believe that this is not a major limitation as the guiding RNA was optimized to limit off-target interactions. BLAST search indicated that one sequence contained 2 mismatches compared to the targeted sequence, but these mismatches were located close to the PAM motif, lowering the chance of off-target mutagenesis[423]. Except for the *XPO1* gene, all other retrieved gene sequences contained at least 3 mismatch mutations when compared to the guiding RNA. Out of the clones that we selected, only some effectively integrated the desired mutations in both alleles of *XPO1* which results in homozygous mutant cells. The remaining clones appeared to be heterozygous for the desired mutations. Moreover, these remaining clones either contained a wild-type allele or a dysfunctional indel containing allele as caused by inefficient NHEJ[224, 225]. This observed distribution is explained as recombination of only one of the *XPO1* alleles was sufficient to confer resistance to SINE compounds. Although homology-directed repair is generally inefficient, we managed to obtain homozygous clones relatively easy. This might be explained by the enrichment of mutated clones by drug-selective pressure in combination with a modified transfection protocol, in which we re-transfected the cells 24 hours after the initial transfection a second time with only repair donor template. We speculated that the single stranded repair donor oligonucleotide is degraded quickly inside the cell. This would result in low concentrations of the oligonucleotide at the time the Cas9 endonuclease protein and the appropriate targeting RNA are adequately expressed and present in the cells after transfection. Thus, we reasoned that additional delivery of the donor repair template after the first transfection might increase the chance of successful recombination.

Finally, the establishment of XPO1 mutant cells may have other important implications. They could form the basis for the identification of new XPO1 inhibitors that do not interact covalently with the cysteine$_{528}$ residue. They also provide a powerful tool to study the role of cysteine$_{528}$ in the function and regulation of XPO1. Cysteines are known to be susceptible to post-translational modification that impact protein function[424-427] and in this context, S-nitrosylation of the cysteine$_{528}$ residue has been suggested to regulate XPO1[428].

Conclusion

In this work we reached the so-called "gold standard" proof for drug-target validation in human cells by introducing a genomic mutation in the *XPO1* gene of cancer cells that confers these cells with resistance to the SINE compound selinexor (KPT-330), the clinical candidate of a class of small molecule inhibitors of nuclear export (SINE). Exportin-1 (XPO1/CRM1) is the key nuclear exporter for a broad range of cargo-proteins involved in cellular transcription, growth regulation and tumour suppression and is considered a therapeutic anticancer target. We applied CRISPR/SpCas9 genome editing in combination with homology directed repair to introduce a C528S mutation in XPO1 of leukaemia cells and succeeded in generating monoclonal hetero- and homozygous mutant cell lines. This C528S substitution confers resistance to selinexor, validating XPO1 as the prime target for selinexor and demonstrating for the first time the unique selectivity of the drug for its target in human cells. Our study also illustrates the general applicability of genome editing in mammalian cells using CRISPR/Cas9 for drug-target validation and target selectivity determination of anticancer drugs. In addition, it shows this technique is feasible for the study of recessive drug-resistance mutations. Finally, the generation of these mutant XPO1 cell lines is quite unique since in higher eukaryotes the XPO1 cysteine$_{528}$ residue is well conserved. The generated cell lines are therefore a powerful tool that opens opportunities to identify non-covalent XPO1 inhibitors and to study the role of the cysteine$_{528}$ residue in the function and regulation of XPO1 in the higher eukaryotic cell.

Methods

Cell Culture

HL-60, Jurkat, HeLa and K-562 cancer cell lines were obtained from the American Type Culture Collection (ATCC) and HAP1 cells, which are derived from KBM-7 chronic myelogenous leukemia cells, were obtained from Horizon Discovery. All cells were cultured in 75cm^2 flasks at 37°C and 4,5% CO_2. Cells were grown in complete RPMI 1640 (Jurkat), IMDM (HL-60, K-562, HAP1) or DMEM (HeLa) medium containing 2 mM L-glutamine and supplemented with 10% fetal bovine serum (FBS) and 20 µg/mL gentamicin. Suspension cells were kept between 3 to 15 x 10^5 cells per mL and passaged three times per week. HAP1 and HeLa cells were passaged three times per week by detachment with 0.05% trypsin in EDTA.

CRISPR/Cas9-mediated knock-in

HL-60, Jurkat, and K-562 cells were transfected using electroporation with a Neon Transfection system (Invitrogen, Life Technologies). Briefly, cells were collected by centrifugation at 400× *g* and resuspended in Resuspension Buffer R (Invitrogen, Life Technologies) according to the instructions. Purified and highly concentrated DNA plasmids expressing SpCas9-NLS$_{SV40}$ (0.5 µg), single guide RNA targeting the Cys528 codon in *XPO1* (0.5 µg) and a single stranded donor oligonucleotide carrying the C528S missense mutation (1 µg) were added to the resuspended cells. Next the mixture, containing 100,000-200,000 cells per 10 µL, was electroporated with the following settings: 10 µL, 1,350V, 10 ms and 3 pulses. Following electroporation, cells were immediately plated in 1 mL of antibiotic free RPMI 1640 or IMDM medium (Gibco, Life Technologies) containing 10% FBS in a 24-well plate and cultured at 37°C in a CO_2 incubator. The medium was refreshed after two days and KPT-185 was added to a final concentration of 100 nM. Cells were then maintained and grown in the presence of the compound over a period of 1-3 weeks following standard cell culture guidelines. After this period, surviving cells were harvested and cultured further. When cultures were sufficiently grown, cells were plated at a density of 0.5 cells/well in 96-well plates in 20% FBS containing medium to obtain single cell derived colonies. Colonies were grown for 2-6 weeks and were regularly screened. Ultimately, colonies were harvested, and the genomic DNA was extracted for Sanger sequencing.

For knock-in of the Glu571Lys missense mutation in XPO1, HAP1 cells were transfected with SpCas9-NLS$_{SV40}$ together with a plasmid encoding an sgRNA-targeting *XPO1* at the codon for Glu571 and a donor repair template ssDNA oligonucleotide (Integrated DNA technologies) carrying the Glu571Lys mutation and 1 additional silent mutation. Cells were electroporated using the Neon Electroporation system (Thermo Fisher Scientific; 10 µL, 1,475V, 10 ms, 3x, buffer R) and plated immediately in 1 mL prewarmed, antibiotic-free IMDM medium supplemented with 10% v/v FBS. Five days after transfection, cells were diluted (0.5 cell/well) into 96-well plates and scored for the formation of single colonies after 3 days. Single colonies were grown for a period of 2 weeks and the DNA was extracted.

DNA constructs

The XPO1-YFP and Rev-BFP plasmids were cloned in house and contained a CMV promotor to express XPO1-YFP or Rev-BFP. The Cas9 expression construct, CMV-SpCas9-NLS$_{SV40}$-HA-linker, and the sgRNA constructs targeting XPO1 at the Cys528 or Glu571 codons were obtained from ToolGen/Labomics. The sgRNAs were expressed from a U6 promotor and contained the following targeting sequence: 5′-GGATTATGTGAACAGAAAAG (C528) or 5′-AAGCTGTTCGAATTCATGCA (E571).

The oligonucleotides used for homologous recombination consisted of 135-base single-stranded oligodeoxynucleotide (ssODN) molecules containing two outer arms (50–60 bp) homologous to the endogenous *XPO1* genomic region flanking the exon coding for C528 or E571 and were synthesized by Integrated DNA Technologies (IDT). The C528 oligonucleotide contained two point mutations at the C528 coding triplet to provide the template for the C528S mutation as well as three silent mutations to prevent SpCas9-mediated cleavage of the mutated allele. It consisted of the following sequence: 5′-GCTAAATAAGTATTATGTTGTTACAATAAATAATACAAATTTGTCTTATTTACAGGATCTATTAGGATTA**TCA**GAACAGAA**gc**G**c**GGCAAAGATAATAAAGCTATTATTGCATCAAATATCATGTACATAGTAGG (lowercase indicates additional silent mutations). The E571 oligonucleotide contained 1 point mutation at the E571 coding triplet together with 1 silent mutation: 5′-GTCAATACCCACGTTTTTTGAGAGCTCACTGGAAATTTCTGAAGACTGTAGTTAACAAGCTGTTC**AAA**TT**t**ATGCATGGTAAATCTCTTTCTTTACTATATTTTGCTTTTATTTTTATTGAAGAAAATAAATGAATGTTTTTGTCT-3′ (lowercase indicates additional silent mutation).

DNA extraction and sequencing

Following cell growth, confluent 96-wells containing single colonies were spinned down, washed and then lysed in Bradley lysis buffer (10 mM Tris-HCl (pH 7.5), 10 mM EDTA, 0.5% SDS, 10 mM NaCl and 1 µg/mL proteinase K) for two hours at 56°C. The genomic DNA was extracted from the lysates using ethanol/salt precipitation coupled to centrifugation. The target site, the DNA sequence around *XPO1*, was amplified by PCR with the following primers: fwd: 5′-TCTGCCTCTCCGTTGCTTTC, rv: 5′-CCAATCATGTACCCCACAGCT. After PCR, the products were sequenced by Sanger Sequencing (Macrogen) using a forward and reverse primer (5′-TGTGTTGGGCAATAGGCTCC, 5′-GGCATTTTTGGGCTATTTTAATGAAA, respectively).

Compounds

KPT-185, selinexor (KPT-330) and KPT-9058 were provided by Karyopharm Therapeutics (Newton, MA). All compounds were dissolved in DMSO.

RNA extraction and quantitative PCR

Total RNA was extracted from Jurkat cells using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. Quantitative RT-PCR was performed on the mRNA using the Express One-Step SuperScript qRT-PCR Kit (Invitrogen) with Prime Time qPCR probes (IDT) against mRNA of *XPO1* exon 9–10 (Hs.PT.58.4000376), mRNA of *XPO1* exon 23–25 (Hs.PT.58.288478), mRNA of *ACTB* exon 1–2 (Hs.PT.39a.22214847), and mRNA of *GAPDH* exon 2–3 (Hs.PT.39a.22214836) on the 7500 Fast Real-Time PCR System (Applied Biosystems). All assays were performed with three biological replicates and three technical replicates, and comparative computed tomography was used to analyse the results with *ACTB* and *GAPDH* serving as internal controls. Statistical significance was determined with an unpaired two-tailed Student's t test.

Cell growth and size assays

Growth rates of mutant and wild-type cell lines were determined with the CyQuant Direct Cell Proliferation kit (ThermoFisher Scientific). Initially a standard curve was obtained to determine the relation of the fluorescent signal to the number of cells present in the sample. For the experiment, 20,000 Jurkat cells or 10,000 K-562 and HL-60 cells were seeded in different 96-well plates on day 0 in complete RPMI or IMDM medium respectively. The cells were grown over a period of 3-4 days and the number of cells was determined every 24 hours by measuring fluorescence of the CyQuant Direct Cell Proliferation reagent with a SaFire II microplate reader according to the manufacturer's instructions. The obtained data points were compared to the previously obtained standard curves and cell numbers were normalized to day 0. Growth curves were then obtained by fitting the data with an exponential growth curve in GraphPad Prism and statistical significance was determined. The experiment was performed in triplicate and repeated at least once. For the determination of average cell sizes, all cells were analysed during 3 different passages with the LUNA automated cell counter (Logos Biosystems). Statistical significance between cell sizes was determined with ANOVA.

Cell Viability assays

Cell viability assays were performed by plating 4,000-20,000 cells in 96-well plates containing DMSO or a dilution of KPT-330 or KPT-9058. Cells were incubated for 72 hours at 37°C and 4.5% CO_2. Cell viability was assessed with the CellTiter 96® AQueous Non-Radioactive Cell Proliferation Assay reagent (Promega) according to the manufacturer's instructions. Absorbance of the samples was measured at 490 nm using a SaFire II microplate reader (Tecan). All assays were performed in duplicate and each experiment was repeated at least three times. The data was analysed using a log-based 4 parameter model in GraphPad Prism and the bottom value for the EC_{50} calculation was set by hand for the mutant K-562 cells.

Apoptosis analysis using flow cytometry

Apoptosis and cell death following 12hr incubation of Jurkat cells with the compounds or DMSO (control) were measured with the AlexaFluor488 Annexin V/Dead Cell Apoptosis Kit (Life Technologies). Briefly 200,000 cells were incubated overnight with test compounds, carrier (DMSO), and/or the apoptosis inhibitor Q-VD-OPh ((3S)-5-(2,6-Difluorophenoxy)-3-[[(2S)-3-methyl-1-oxo-2-[(2-quinolinylcarbonyl)amino]butyl]amino]-4-oxo-pentanoic acid hydrate; Sigma-Aldrich). The following morning, cells were incubated with AlexaFluor488 conjugated Annexin V and propidium iodide (PI) according to the manufacturer's instructions. Samples were then analysed using the green and red fluorescence with a FACSCanto II flow cytometer (BD Biosciences) and the FACSDiva software (BD Biosciences).

Cell cycle analysis using flow cytometry

Cell cycle analysis on Jurkat cells was performed using the BD Cyclotest PLUS (BD Biosciences) kit. Cells were incubated for 24hr with test compounds and harvested using standard protocol. Following harvesting, the cells were treated according to the manufacturer's instructions, and the DNA content was determined by measuring PI fluorescence with the FACSCanto II flow cytometer and FACSDiva software. For both experiments, averages and SDs were calculated using the data of each individual experiment, and significance was determined using an unpaired two-tailed Student's t test considering the F-values for equal or unequal variance.

Classical Western blot analysis

Classical Western blotting was performed according to standard protocol. In short, Jurkat cells were seeded at 0.7×10^6 cells/mL in 24-well plates and incubated with DMSO or compound for 16 hr. Cells were then collected at 400 × g and washed with ice-cold PBS. Supernatant was removed, and cells were lysed in lysis buffer (50 mM Tris-HCl, 50 mM NaCl, 5 mM MgCl2, 0.5% Triton X-100, and 1X halt protease inhibitor; Thermo Fisher Scientific) for 1hr on ice (20×10^6 cells/mL). Whole-cell lysates were then cleared by brief centrifugation. Protein lysates were denatured and subsequently resolved by SDS-PAGE and transferred to a Hybond-P polyvinylidene fluoride (PVDF) membrane (GE Healthcare). Membranes were incubated for 1hr at room temperature in blocking buffer (8% w/v non-fat dry milk in PBS containing 0.05% v/v Tween-20) and thereafter incubated for 12hr at 4°C in PBS containing 0.05% v/v Tween-20 with primary antibodies raised against poly [ADP-ribose] polymerase 1 (PARP-1) (sc-8007; Santa Cruz Biotechnology), caspase-3 (sc-271028; Santa Cruz Biotechnology), CRM1/XPO1 (sc-74454; Santa Cruz Biotechnology), or β-tubulin (ab6046; Abcam). After washing, the membranes were incubated with goat anti-mouse or anti-rabbit IgG horseradish peroxidase-conjugated secondary antibody (sc-2005; sc-2004; Santa Cruz Biotechnology) in PBS containing 0.05% v/v Tween-20 for 20 min at room temperature. Subsequently, membranes were washed, and proteins were detected by chemiluminescence on a Chemidoc MP system (Bio-Rad) after the addition of a luminol/peroxide solution.

Immunofluorescence staining and confocal fluorescence microscopy

Cells were treated with compound or carrier for 4–16hr. Cells were harvested at 400× g, washed in PBS, and transferred into an eight-well chambered Nunc Lab-Tek Coverglass, which had been pre-treated with 0.1% (w/v) poly-l-lysine (Sigma). After treatment, cells were carefully washed with PBS. Cells were then fixed with 4% paraformaldehyde (PFA), washed, and permeabilized. Further treatment was performed according to standard procedures, and cell nuclei were counterstained with DAPI. Employed antibodies included rabbit anti-human RanBP1 (Ab97659; Abcam), rabbit anti-human CRM1/XPO1 (sc-5595; Santa Cruz Biotechnology), and goat anti-rabbit IgG conjugated to Alexa Fluor 488 (A11008; Invitrogen). All images were collected with a Leica TCS SP5 confocal microscope employing a HCX PL APO 63× (NA 1.2)/water immersion objective. Alexa Fluor 488 was detected using the excitation line of 488 nm (Argon Laser), and DAPI was detected using the excitation line of 405 nm (pulsed diode laser). Blue emission was detected between 410 and 480 nm, and green emission was detected between 493 and 565 nm.

For quantification of the cellular XPO1 protein content, wild-type or $XPO1^{C528S}$ Jurkat cells were cultured in four different wells of a chambered coverglass in a randomized manner and then stained for XPO1 as described above. Each well was imaged three times at different positions in a random order. The experiment was repeated two times on different dates. The fluorescence in the green channel was quantified on a per-cell basis employing the Imaris v.7.6.5 image analysis software (Bitplane). Mean pixel intensity per cell values were normalized to compare different experiments. For wild-type and $XPO1^{C528S}$ cells, 2,269 and 1,866 cells were analysed, respectively. Statistical significance was determined with an unpaired two-tailed

Student's t test. For the colocalization experiments, HeLa cells were transfected with the respective plasmids and monitored 1 day after transfection using the above described confocal microscope setup. XPO1 and Rev were visualized by imaging the BFP and YFP tags using 405 nm (BFP) or 488 nm (YFP) excitation lines. Emission was detected between 410 and 480 nm and 505 and 600 nm, respectively.

High content imaging of RanBP1

Jurkat cells were plated in a 96-well imaging plate (Falcon) and treated with a dilution series of selinexor (KPT-330) or carrier (DMSO) for 3 hours. After treatment, cells were washed with PBS and fixed for 10 minutes in 4% PFA. The cells were then washed and remaining PFA was inactivated with PBS containing 0.1M glycine. Cells were permeabilized with PBS containing 0.1% Triton X-100 for 10 minutes. Subsequently, the cells were washed 3 times with PBS and then blocked at 37°C for 2 hours in 10% Normal Goat Serum. The primary anti-human RanBP1 (Ab97659, Abcam) antibody was then added and incubated for 1 hour. After primary staining, the cells were extensively washed and the secondary antibody, Alexa Fluor 488 (A11008, Invitrogen), was incubated for 1 hour in the dark. Finally, cells were washed, and nuclei were counterstained with DAPI. The plate was imaged using the green and blue channels of an ArrayScan XTI High Content Reader (ThermoFisher Scientific). Nuclear and cytoplasmic compartments were segmented, and their average pixel intensities were quantitated with the HCS Studio software. The obtained nuclear/cytoplasmic ratios were analysed with GraphPad Prism. The experiment was repeated two times.

Visualization of XPO1 degradation by SINE compounds

For detection of KPT-330-induced degradation of XPO1, 2×10^6 Jurkat cells were plated in 2 mL supplemented RPMI in 6-well plates and treated overnight. The next morning cells were collected by centrifugation (400× g) and washed in ice-cold PBS. Cell pellets were lysed on ice in RIPA buffer supplemented with 1x HALT protease inhibitors (ThermoFisher Scientific) and cleared from debris by centrifugation at 20,000× g for 10 min at 4°C. The supernatant was prepared according to the manufacturer's instructions (Protein Simple). Proteins were then separated by size (12-230 kDa) and visualized on a Wes system (Protein Simple) with an anti-rabbit HRP conjugated antibody detecting the primary XPO1 (1/12,500, NB100-79802) and β-tubulin (1/3,000, NB600-936) rabbit antibodies. Protein signals were visualized and quantified with the Compass software, v2.7.1 (Protein Simple).

XPO1 pull down with KPT-9058

For pull down of XPO1 using KPT-9058, 5×10^6 Jurkat cells were plated in 1.5 mL of supplemented RPMI inside 6-well plates. After a few hours inside the incubator, 1.5 mL of RPMI containing 2 μM of KPT-9058 was added to the cells (final concentration 1 μM). Cells were incubated for 2 hours and then collected by centrifugation (400× g) and washed in ice-cold PBS. Cell pellets were lysed on ice in RIPA buffer supplemented with 1x HALT protease inhibitors (ThermoFisher Scientific) and cleared from debris by centrifugation at 20,000× g for 10 min at 4°C. Protein concentrations were measured using a colorimetric BSA protein assay (Pierce). A fraction of the lysis mixture was taken for quantification of β-tubulin and XPO1 total protein. The remaining extracts were allowed to bind to Dynabeads MyOne Streptavidin T1 (Life Technologies) by rotating overnight at 4°C to capture KPT-9058. Following overnight incubation, the beads were washed 5 times in modified RIPA buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 1% NP-40 (IGEPAL CA-630), 0.1% sodium deoxycholate, 1 mM EDTA) and subsequently boiled for 10 minutes in 0.5% SDS containing 1x sample buffer (Protein Simple) to elute captured proteins. Proteins were separated by size (12-230 kDa) and visualized on a Wes system (Protein Simple) as described above.

Acknowledgements
We thank Lotte Bral, Astrid d'Hoore, and Bob Massant for their exceptional technical help and support.

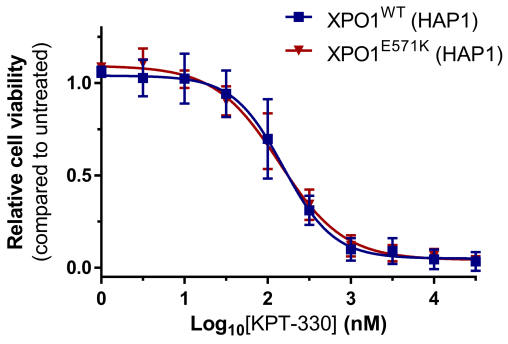
Competing interests
E.B., S.S., M.C. and Y.L. are employees of Karyopharm Therapeutics. D.D. has a license arrangement on SINE XPO1 inhibitors. The remaining authors have no potential conflict of interest to declare.

Author contributions
J.E.N., M.J., E.V., and T.V. performed experiments and analyzed the data. J.E.N., T.V., E.V., M.J., and D.D. designed and discussed experiment setups. E.B., S.S., M.C., and Y.L. provided essential reagents and input. D.D. designed the study. J.E.N. designed the figures and J.E.N. and D.D. cowrote the manuscript with input from all authors.

Supplementary information

	Allele 1 (variable allele)	Allele 2 (mutant allele)
Clone 1 (Heterozygote)	…taggatta**tgt**gaacagaa...GCgCggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 2 (Heterozygote)	…taggatta**tgt**gaacagaa...GCgCggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 3 (Heterozygote)	…taggatta**tgt**gaacagaa...aagaggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 4 (Heterozygote)	…taggatta**tgt**gaacagaa...aagaggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 5 (Homozygote)	…taggatta**tCA**gaacagaa...GCgCggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 6 (Homozygote)	…taggatta**tCA**gaacagaa...GCgCggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 7 (NHEJ deletion)	…taggatta**tgt**gaa-----...---aggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 8 (NHEJ insertion)	…taggatta**tgt**gaacagaa***GCT***aagaggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 9 (NHEJ insertion)	…taggatta**tgt**gaacaga***GGCA***aagaggca…	…taggatta**tCA**gaacagaaGCgCggca…
Clone 10 (NHEJ insertion)	…taggatta**tgt**gaacagaa***CCG***aagaggca…	…taggatta**tCA**gaacagaaGCgCggca…

Supplementary Table 1: *XPO1 sequence around the cys$_{528}$ codon of resistant Jurkat clones.*
The encoding bases for the cysteine$_{528}$ (tgt) or serine$_{528}$ (tca) residue of *XPO1* are shown in bold. Inserted or mutated bases are indicated by italic script and additional silent mutations are marked by magenta coloring. All cells were transfected with CRISPR/SpCas9, an sgRNA targeting *XPO1* at codon C528 and a repair donor template. After transfection, cells were selected with KPT-185.



Supplementary Figure 1: *the E571K mutation does not alter sensitivity to KPT-330.*
Cellular viability 72 hours after treatment of XPO1 wild-type and CRISPR/Cas9-edited XPO1^{E571K} mutant HAP1 cells with increasing concentrations of KPT-330. Data points represent averages ± s.d. obtained from three experiments performed in duplicate. No statistical difference between the dose-response curves was detected (ANOVA).

CHAPTER IV

Article 2: drug target deconvolution with CRISPR/Cas mutagenesis scanning

Target identification of small molecules using large-scale CRISPR-Cas mutagenesis scanning of essential genes

Jasper Edgar Neggers[1], Bert Kwanten[1], Tim Dierckx[2], Hiroki Noguchi [3], Arnout Voet[3], Lotte Bral[1], Kristien Minner[1], Bob Massant[1], Nicolas Kint[4], Michel Delforge[4], Thomas Vercruysse[1], Erkan Baloglu[5], William Senapedis[5], Maarten Jacquemyn[1] and Dirk Daelemans[1]

[1]KU Leuven Department of Microbiology and Immunology, Laboratory of Virology and Chemotherapy, Rega Institute for Medical Research, Herestraat 49, 3000 Leuven, Belgium.

[2]KU Leuven Department of Microbiology and Immunology, Laboratory of Clinical and Epidemiological Virology, Rega Institute for Medical Research, Herestraat 49, 3000 Leuven, Belgium.

[3]KU Leuven Department of Chemistry, Biochemistry Molecular and Structural Biology Section, Celestijnenlaan 200 G, 3001 Leuven, Belgium.

[4]Department of Hematology, University Hospital Leuven, Herestraat 49, 3000 Leuven, Belgium.

[5]Karyopharm Therapeutics Inc, 85 Wells Ave, Newton, MA 02459, USA.

Abstract

Unraveling the mechanism of action and molecular target of small molecules remains a major challenge in drug discovery. While many cancer drugs target genetic vulnerabilities, loss-of function screens fail to identify essential genes involved in drug mechanism of action. Here, we report CRISPRres, a CRISPR-Cas-based genetic screening approach to rapidly derive and identify drug resistance mutations in essential genes. It exploits the local genetic variation created by CRISPR-Cas-induced non-homologous end-joining (NHEJ) repair to generate a wide variety of functional in-frame mutations. Using large sgRNA tiling libraries and known drug–target pairs, we validate it as a target identification approach. We then apply CRISPRres to the anticancer agent KPT-9274 and identify nicotinamide phosphoribosyltransferase (NAMPT) as its main target. These results present a powerful and simple genetic approach to create many protein variants that, in combination with positive selection, can be applied to reveal the cellular target of small-molecule inhibitors.

Adapted from the original publication:
Neggers et al. *Nature Communications,* volume 9, article number 502, February 5 (2018)
doi:10.1038/s41467-017-02349-8

Introduction

Identifying the cellular target of a chemical hit with valuable activity is a crucial step in drug discovery and development[36]. However, unraveling the molecular target of a small molecule drug remains a challenging, laborious and complex process. Although many target deconvolution methods[39, 44], such as chemical proteomics, have successfully been applied, they often reveal more than one plausible candidate target protein and carry the risk of identifying interactions that are not related to the compound's activity. The gold standard proof for a drug's target is the identification of functional mutations that confer resistance in a cellular context. For this reason, genetic screens in particular, are very powerful tools for drug mechanism of action studies[38]. However, current screens either are not well suited to identify essential genes or require whole exome sequencing combined with complex bio-informatics to deconvolute the relevant drug resistance conferring mutations. For example, loss-of-function approaches have been applied to obtain drug resistance[214, 222, 254, 429], but innately lack the ability to comprehensively detect gain-of-function mutations and fail to nominate essential proteins involved in drug mechanism of action. Classical step-wise drug resistance selection enables selection of gain-of-function mutations but is laborious[202] and often results in off-target multi-drug resistance[201]. Recently, chemical mutagenesis to increase the occurrence of single-nucleotide variants has been described[205]. However, up until now, this chemical mutagenesis approach has only been applied to identify loss-of-function mutations conferring resistance to the prototype acute myeloid leukemia drug 6-thioguanine. It remains to be investigated whether this approach can also detect gain-of-function resistance mutations. Another bottleneck of these general random mutagenesis approaches is the discovery of the underlying resistance mutations. They require sequencing of the large human exome in individual clones[120, 121, 203, 205] while the genomic heterogeneity of the cell line makes the deconvolution of the relevant resistance conferring mutations especially challenging. As such, the field would greatly benefit from an approach that can accelerate the drug resistance selection process and simplify subsequent identification of the relevant drug resistance mutations. Moreover, because many cancer drugs target essential proteins there is a strong need for a method that can easily generate and identify drug resistance mutations in essential genes.

Drawing a parallel to the use of UV-mediated double strand breaks (DSBs) to enhance mutagenesis[430], we reasoned that introduction of DSBs by targeted endonucleases, such as SpCas9, and the subsequent error-prone repair via non-homologous end-joining (NHEJ) may be exploited for rational protein mutagenesis to facilitate drug resistance selection. Here, we describe a CRISPR-based method, entitled CRISPR-induced **r**esistance in **e**ssential gene**s** (CRISPRres), to rapidly acquire and identify drug resistance mutations. We show that large-scale CRISPR sgRNA gene tiling libraries can be applied as a genetic screening approach in cancer cells to identify the molecular target of a chemical inhibitor. Finally, we also demonstrate the methodology is compatible with the Class 2 type V AsCpf1 CRISPR system, increasing the resolution of the method.

Results

Rapid generation of drug resistant variants with CRISPR-Cas9

To develop the method, we first designed single sgRNAs targeting known resistance hotspots in genes sensitive to three cancer drugs: KPT-185, a preclinical analogue of the XPO1 inhibitor selinexor[396, 397, 399, 420], ispinesib, an antineoplastic kinesin-5 (KIF11) inhibitor[120, 431] and triptolide, an anti-proliferative agent targeting ERCC3[121, 432] (**Fig. 1a** and **b**). The respective sgRNAs were expressed together with SpCas9 in the chronic myeloid leukemia derived HAP1 cell line and treated with 4 different concentrations of the corresponding drug. Within a few days of treatment, colonies that were resistant to the drugs appeared on the culture plates (**Fig. 1c** and **d**). Next-generation sequencing of the targeted hotspot loci of these resistant colonies revealed known as well as many novel resistant protein variants (**Fig. 1e** and **Supplementary Fig. 1a, 2** and **3**). Mutations were mainly localized within 17bp upstream of the SpCas9 cleavage site and consisted of insertions, deletions and missense mutations (**Fig. 1f** and **g**). Most of the sequences consisted of in-frame mutations, but some frameshift and nonsense mutations were also detected. Because the targeted genes are essential for survival, this suggests some of the cells had turned diploid during the experiment, a phenomenon known to occur spontaneously in HAP1 cells[433]. For XPO1, more than 40 different in-frame variants containing a mutation or deletion of the cysteine$_{528}$ residue were detected (**Fig. 1e** left panel and **Supplementary Fig. 1a**). This agrees with the fact that this single cysteine residue is absolutely essential for KPT-185/selinexor activity[397, 399, 420]. However, up until now, only a single serine substitution had been reported to confer cancer cells with resistance to both drugs[420]. As we were surprised to see an essential protein such as XPO1 accommodate such a wide variety of diverse mutations in a functional domain, we derived different single-cell clones from the original compound selected cell pools to confirm the resistance-conferring mutations by Sanger sequencing and cell viability in the presence of drug (**Supplementary Fig. 1b** and **c**).

For KIF11, the majority of the sequencing reads contained a 9 base pair deletion from codon D130 to L132 (**Fig. 1e** middle panel and **Supplementary Fig. 2**). Interestingly, the majority of sequences shared a deletion or mutation of L132, which appears to be sufficient for drug resistance per se. Mutation of residue A133, known to confer ispinesib resistance[120], was not readily detected, which might be explained by the fact that the codon for A133 is at the 3' of the SpCas9 cut site. Indeed, mutations were mainly observed upstream of the SpCas9 cut site, 5' from the NGG protospacer adjacent motif (PAM) (**Fig. 1g** middle panel). Interestingly, in contrast to the selinexor resistance mutations, almost all ispinesib resistant sequence alterations were deletions. For ERCC3, we mainly identified in-frame insertions and some deletions between codons K165 and K167 (**Fig. 1e** right panel and **Supplementary Fig. 3**). We also targeted an additional resistance hotspot at ERCC3 codon D54, which resulted in in-frame deletions of 1-8 amino acids consistent with previous observations[121] (**Supplementary Fig. 4**). To confirm that drug resistance was obtained by targeted CRISPR-Cas9-mediated mutagenesis and not by multidrug resistance, the drug-resistant colonies were treated with all three drugs and no cross-resistance between the three drug resistant cells pools was observed (**Supplementary Fig. 5**). To demonstrate this mutagenesis methodology can be broadly applied to other cell types, similar results were obtained in the pseudodiploid acute pro-myelocytic leukemia HL-60 and colon cancer HCT 116 cell lines (**Supplementary Fig. 6** and **7**).

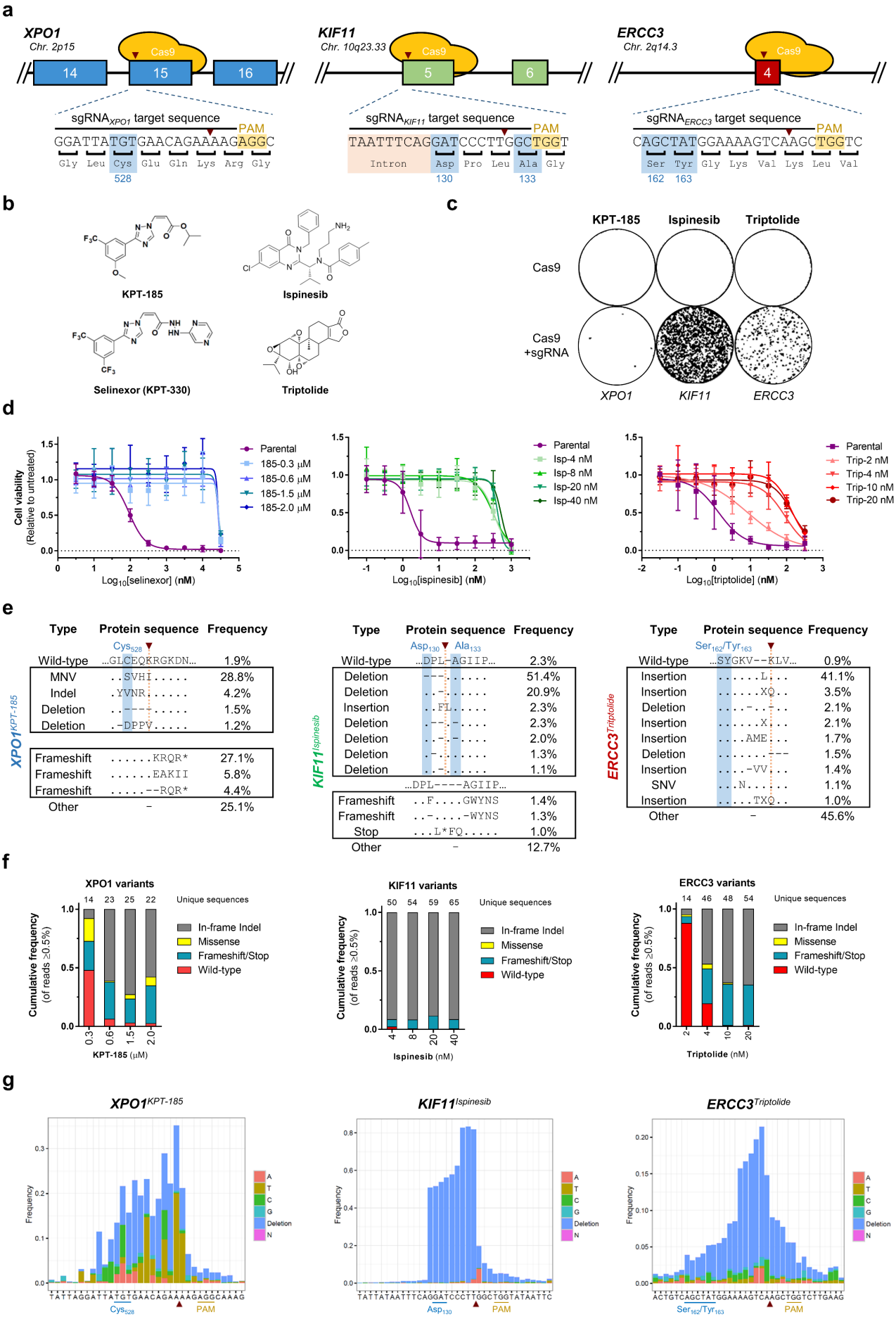

a
XPO1
Chr. 2p15
Cas9
14
15
16
sgRNA$_{XPO1}$ target sequence
PAM
GGATTATGTGAACAGAAAAGAGGC
Gly Leu Cys Glu Gln Lys Arg Gly
528
KIF11
Chr. 10q23.33
5
6
sgRNA$_{KIF11}$ target sequence
TAATTTCAGGATCCCTTGGCTGGT
Intron Asp Pro Leu Ala Gly
130 133
ERCC3
Chr. 2q14.3
4
sgRNA$_{ERCC3}$ target sequence
CAGCTATGGAAAAGTCAAGCTGGTC
Ser Tyr Gly Lys Val Lys Leu Val
162 163
b
KPT-185
Ispinesib
Selinexor (KPT-330)
Triptolide
c
KPT-185
Ispinesib
Triptolide
Cas9
Cas9 +sgRNA
XPO1
KIF11
ERCC3
d
Cell viability (Relative to untreated)
Log$_{10}$[selinexor] (nM)
Parental
185-0.3 µM
185-0.6 µM
185-1.5 µM
185-2.0 µM
Log$_{10}$[ispinesib] (nM)
Isp-4 nM
Isp-8 nM
Isp-20 nM
Isp-40 nM
Log$_{10}$[triptolide] (nM)
Trip-2 nM
Trip-4 nM
Trip-10 nM
Trip-20 nM
e
XPO1KPT-185
Type Protein sequence Frequency
Cys528
Wild-type ...GLCEQKRGKDN... 1.9%
MNV ..SVHI..... 28.8%
Indel .YVNR..... 4.2%
Deletion ..-----..... 1.5%
Deletion .-DPPV..... 1.2%
Frameshift .....KRQR* 27.1%
Frameshift .....EAKII 5.8%
Frameshift .....--RQR* 4.4%
Other - 25.1%
KIF11Ispinesib
Type Protein sequence Frequency
Asp130 Ala133
Wild-type ...DPL-AGIIP... 2.3%
Deletion ---..... 51.4%
Deletion ...-..... 20.9%
Insertion ..FL..... 2.3%
Deletion .--.-.... 2.3%
Deletion ..-.-.... 2.0%
Deletion .-..-..... 1.3%
Deletion ..-..... 1.1%
...DPL----AGIIP...
Frameshift ..F....GWYNS 1.4%
Frameshift ..-....-WYNS 1.3%
Stop ...L*FQ..... 1.0%
Other - 12.7%
ERCC3Triptolide
Type Protein sequence Frequency
Ser162/Tyr163
Wild-type ...SYGKV--KLV... 0.9%
Insertion .....L... 41.1%
Insertion .....XQ.. 3.5%
Deletion ....-..-.. 2.1%
Insertion .....X... 2.1%
Insertion ....AME... 1.7%
Deletion .....--- 1.5%
Insertion ....-VV... 1.4%
SNV ...N..... 1.1%
Insertion .....TXQ.. 1.0%
Other - 45.6%
f
XPO1 variants
14 23 25 22
Unique sequences
In-frame Indel
Missense
Frameshift/Stop
Wild-type
Cumulative frequency (of reads ≥0.5%)
KPT-185 (µM)
KIF11 variants
50 54 59 65
Ispinesib (nM)
ERCC3 variants
14 46 48 54
Triptolide (nM)
g
XPO1KPT-185
Frequency
A
T
C
G
Deletion
N
TATTAGGATTATGTGAACAGAAAAGAGGCAAAG
Cys528
PAM
KIF11Ispinesib
TATTATAATTTCAGGATCCCTTGGCTGGTATAATTC
Asp130
ERCC3Triptolide
ACTGTCAGCTATGGAAAAGTCAAGCTGGTCTTGAAG
Ser162/Tyr163

Figure 1: *CRISPR-Cas9-induced NHEJ repair facilitates rapid selection of drug resistant variants.*
a. Representation of sgRNAs targeting resistance hotspots for KPT-185/selinexor (*XPO1* codon C528), ispinesib (*KIF11* codons D130 and A133), and triptolide (*ERCC3* codons S162 and Y163). The SpCas9 cleavage site is indicated (red arrowhead) and residues conferring drug resistance when altered are highlighted in blue.
b. Chemical structures of the used antineoplastic agents.
c. SpCas9-induced NHEJ repair at resistance hotspots facilitates rapid selection of resistant colonies. Cells were transfected with SpCas9 plasmid (top) or co-transfected with SpCas9 and sgRNA expressing plasmids (bottom). KPT-185 (300 nM), ispinesib (4 nM), or triptolide (10 nM) were added 48hr after transfection and selection was maintained for 7–10 days before visualization. The experiment was performed with four different compound concentrations (KPT-185: 0.3, 0.6, 1.5, 2 μM; ispinesib: 4, 8, 20, 40 nM; triptolide: 2, 4, 10, 20 nM) and replicated at least once (see also **Supplementary Figs. 1, 2,** and **3**), but only one representative concentration is shown.
d. Cell viability assays of wild-type (parental) and the different resistant cell populations obtained after selection of mutagenized cells with four different concentrations of compound. Data points are normalized relative to untreated cells and represent mean ± s.d. obtained from three experiments performed in triplicate.
e. Amino-acid sequence variants present in the resistant cell pools described in panel **c** as determined by targeted amplicon sequencing analysis with CrispRVariants. The wild-type sequence is shown for reference and resistance hotspot residues are highlighted in blue. The SpCas9 cleavage site is indicated (red arrowhead). For a complete list of amino-acid variants from all resistant cell pools and replicate experiments see **Supplementary Figs. 1a, 2,** and **3**.
f. The relative abundance of all alleles with a read frequency ≥0.5% is shown and categorized per sample into four different mutation types. The values represent averages of two (*XPO1*, *ERCC3*) or three (*KIF11*) experiments for each concentration of compound.
g. Single-nucleotide variant occurrence at the sgRNA target sites observed across all mutagenized and selected cell pools from **c**. The red arrowhead indicates the Cas9 cut site. The values represent averages of eight (*XPO1*, *ERCC3*) or 12 (*KIF11*) experiments.

Validation of drug resistant variants

To validate the above observed drug resistance mutations, we reinstalled one of the major protein variants ($XPO1^{C528S,\ E529V,\ Q530H,\ K531I}$; $KIF11^{L132\Delta}$; $ERCC3^{V166\_K167insL}$) into their native locus in parental cells using CRISPR-Cas9-mediated homology-directed repair (HDR) of an ssDNA repair donor template (**Fig. 2a**). Silent mutations not identified in our original selection were included to control for HDR (**Fig. 2b**). All three donor templates provided strong protection against respective drug treatment (**Fig. 2a** and **c**), validating the results obtained by the targeted CRISPR-Cas9-induced mutagenesis without HDR template. We also assessed protein function of the resistant variants and how this is affected by the respective drugs. For this purpose, we visualized the XPO1-dependent nuclear export of a GFP protein fused to the nuclear localization signal (NLS) of SV40 and the nuclear export signal (NES) from the prototype XPO1 cargo protein kinase inhibitor (PKI). This fusion protein actively shuttles between the nucleus and cytoplasm but localized in an XPO1-dependent manner to the cytoplasm in steady state (**Fig. 2d** panel a)[397]. CRISPR-Cas9-mediated knockout of XPO1 in both wild-type and mutant cells resulted in nuclear accumulation of the reporter protein, demonstrating functionality of the resistant mutant XPO1 protein (**Fig. 2d** panel b). Upon inhibition of the XPO1-mediated nuclear export by selinexor, the GFP fusion protein accumulated in the nucleus of wild-type cells, while in $XPO1^{C528S,\ E529V,\ Q530H,\ K531I}$ mutant cells its localization remained unaffected (**Fig. 2d** panel c), demonstrating the mutation confers resistance to inhibition of XPO1 function by selinexor.

As an additional confirmation of drug resistance, we further demonstrated the XPO1 resistance mutations prevent the drug from binding directly to XPO1 by pull-down of XPO1 protein using a biotinylated analogue of selinexor (**Fig. 2e**). For the KIF11$^{L132\Delta}$ resistance mutation, the formation of mitotic spindles in genome-edited KIF11$^{L132\Delta}$ cells was like that in wild-type cells (**Fig. 2f** panel a). Knockout of *KIF11* disrupted bipolar spindle formation in both wild-type and KIF11$^{L132\Delta}$ mutant cells (**Fig. 2f** panel b), demonstrating the mutant protein is functional. Furthermore, the KIF11$^{L132\Delta}$ mutation protected cells from monopolar mitotic spindle formation induced by ispinesib, this in contrast to treated wild-type cells (**Fig. 2f** panel c). Taken together these results demonstrate that spontaneous genetic variation generated during NHEJ repair at the locus of CRISPR-SpCas9-mediated DSBs in essential proteins can be exploited to significantly accelerate the selection of functional drug resistance mutations; a finding independently confirmed by Donovan et al. and Ipsaro et al.[434, 435].

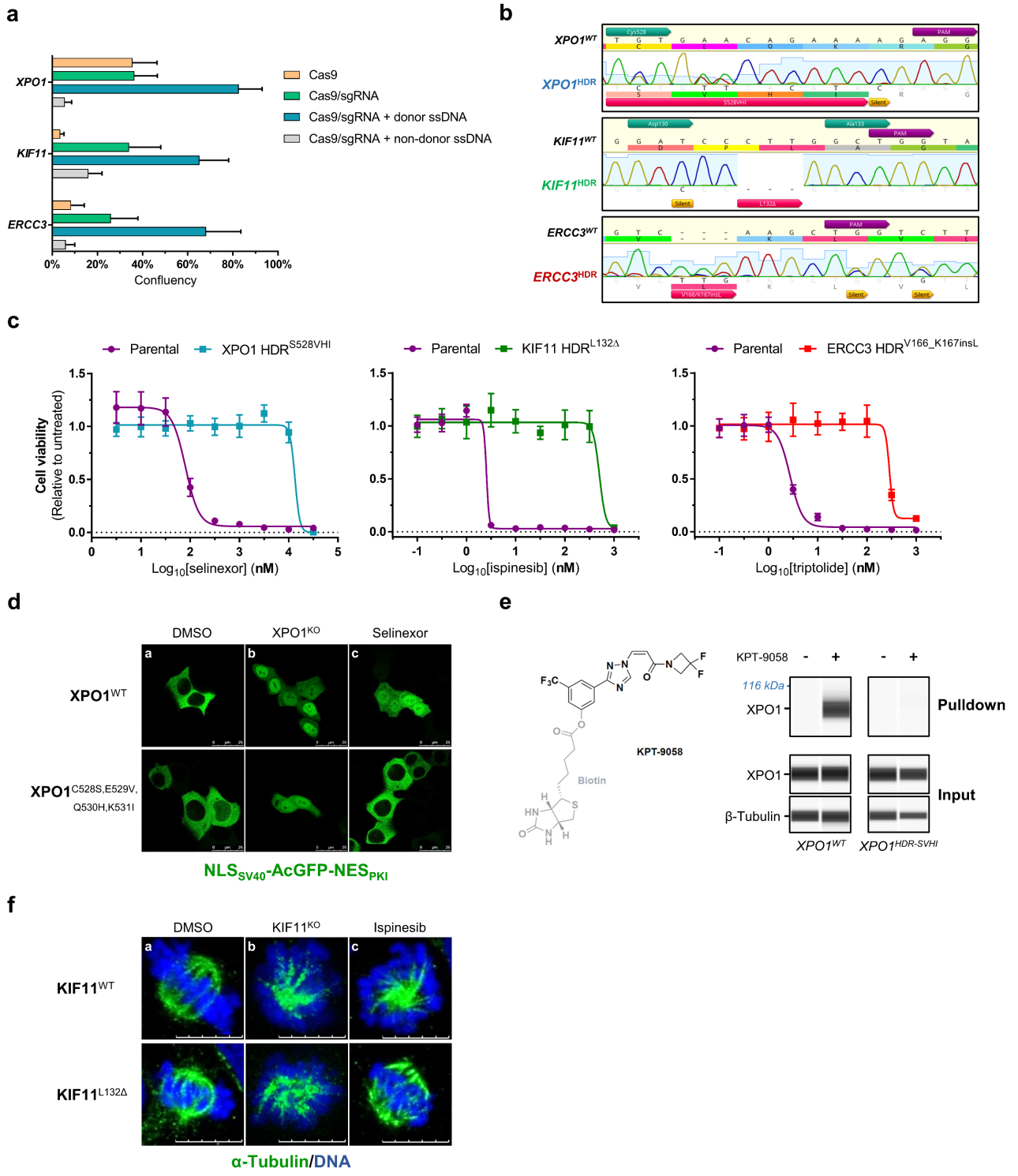


Figure 2: *validation of drug resistance mutations by SpCas9-induced homology-directed repair.*
a. Viability of HAP1 cells transfected with either SpCas9 alone, SpCas9 together with the resistant hotspot targeting sgRNA, Cas9 with the sgRNA and an ssDNA donor template, or Cas9 with the sgRNA and a non-donor ssDNA template after treatment for 5 days with selinexor (2 μM, *XPO1*), ispinesib (10 nM, *KIF11*) or triptolide (40 nM, *ERCC3*), respectively. Data represent the average confluency with standard deviation of a single well as visualized by 16 images taken of the same well by an IncuCyte ZOOM. Cells were treated 2 days after transfection.
b. Sequencing chromatograms of HDR-edited HAP1 cells after treatment. Silent mutations were incorporated to control for HDR.
c. Cell viability assays showing the effect of selinexor, ispinesib or triptolide on wild-type (parental) and HDR-edited HAP1 cells after 72hr. Data points are normalized to untreated cells and represent averages ± s.d. obtained from two experiments performed in triplicate.
d. Confocal fluorescence imaging of parental and HDR-edited *XPO1* HAP1 cells transfected with the NLS-AcGFP-NES XPO1 cargo (green). In addition, some cells were cotransfected with SpCas9 and an sgRNA pool targeting *XPO1* to knockout (KO) *XPO1* (panel b). Two days after transfection, cells were treated with DMSO (panel a) or 2 μM selinexor (panel c) for 3 hours and then imaged. Scale bars: 25 μm.
e. Pulldown of XPO1 from wild-type and mutant XPO1$^{\text{HDR-S528VHI}}$ HAP1 cells. Cells (2.5 x 10^6) were collected and lysed in RIPA buffer after treatment with DMSO or 2 μM KPT-9058 for 3 hours. KPT-9058 was extracted overnight at 4°C by streptavidin affinity chromatography (Dynabeads MyOne Streptavidin T1). The total lysate (input) and extracted fractions (pulldown) were then separated on a Simple Wes system. Indicated proteins were detected by chemiluminescent immunoblotting using primary and secondary-HRP conjugated antibodies.
f. Confocal immunofluorescence of mitotic spindle formation in parental or HDR-edited *KIF11* HAP1 cells treated with DMSO (panel a) or 50 nM ispinesib (panel c) for 2 hours. Some cells were transfected with SpCas9 and an sgRNA pool targeting *KIF11* to knockout *KIF11*(panel b). Scale bars: 7.5 μm.

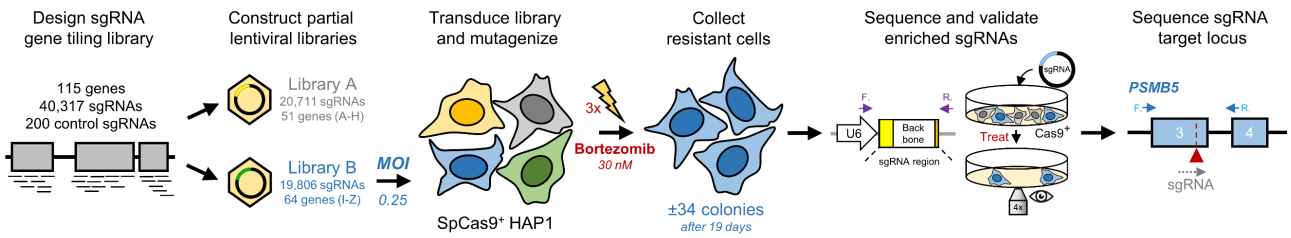
Target identification by sgRNA tiling of multiple genes

To further investigate whether this approach can be applied as a genetic screen to directly identify the molecular target of a chemical inhibitor, we designed a lentiviral tiling sgRNA library spanning all possible NGG PAM sites in the exons from 9 different genes (2,209 sgRNAs), including kinesin-5 (*KIF11*) (**Supplementary Data 1**). HAP1 cells stably expressing SpCas9 were transduced with this lentiviral sgRNA library and subsequently treated with ispinesib for a period of 14 days (**Supplementary Fig. 8a**). Drug resistant colonies formed rapidly and were resistant to ispinesib (**Supplementary Fig. 8b**). Four sgRNAs, which all targeted *KIF11*, were highly enriched in the pool of resistant clones (**Supplementary Fig. 8c** and **d**, and **Supplementary Data 2**). When transfected separately, these sgRNAs were able to confer HAP1 cells with ispinesib resistance, with the sgRNAs targeting codon A133 being the most efficient (**Supplementary Fig. 8e**). The underlying resistance mutations at the genomic locus targeted by this sgRNA consisted of multiple in-frame variations around the KIF11 A133 codon (**Supplementary Fig. 8f**), consistent with previous results (**Fig. 1e** middle panel), and map to the ispinesib-binding site. These results show CRISPR-Cas9 mutagenesis scanning can be used to fish-out the target protein of a drug and at the same time to identify the drug's interaction site from a pool of genes.

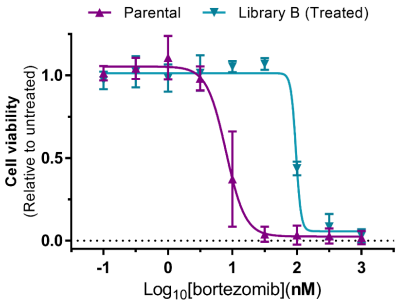
To extend the use of this approach to drugs for which resistance has not yet been developed or for which the target is not known, we designed two complementary lentiviral sgRNA tiling libraries containing the target genes of FDA approved (115 genes, divided in two subpools A and B of each ±20,000 sgRNAs) or investigational antineoplastic drugs (75 genes, divided in two subpools C and D containing ±12,000 sgRNAs each) (**Supplementary Data 3** and **4**). As validation of these libraries, we applied sublibrary B to SpCas9$^+$ HAP1 cells to obtain resistance against the proteasome inhibitor bortezomib, which targets proteasome subunit beta type-5 (PSMB5) (**Fig. 3a**). Surviving colonies appeared and were harvested within 19 days after treatment. These were resistant to bortezomib treatment (**Fig. 3b**) and documentation of

the sgRNAs present in the resistant cell population revealed four highly enriched sgRNAs targeting known PSMB5 resistance hotspots amongst sgRNAs targeting other genes (**Fig. 3c**, **Supplementary Data 5**). Validation of the >100-fold enriched sgRNAs revealed that *PSMB5* was the sole gene that, when mutagenized, conferred bortezomib resistance (**Fig. 3d**). Deep sequencing of the genomic region targeted by these sgRNAs identified multiple in-frame mutations in *PSMB5* (**Fig. 3e**). Most of these mutations are similar to known drug resistance mutations and map to the bortezomib binding site (**Fig. 3f**)[303, 436].

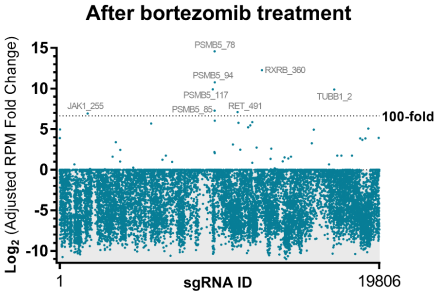
a



b

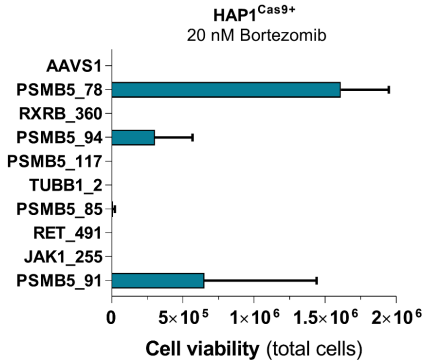


c



sgRNA ID	*Cut Site*	*FC*	*%*
PSMB5_78	***PSMB5*^{A79}**	**24,881**	**12.1%**
RXRB_360	*RXRB*^{Q18}	4,955	23.7%
PSMB5_94	***PSMB5*^{A108}**	**1,761**	**23.6%**
PSMB5_117	***PSMB5*^{Intron1}**	**959**	**22.4%**
TUBB1_2	*TUBB1*^{Q11}	947	0.3%
PSMB5_85	***PSMB5*^{T80}**	**156**	**0.3%**
RET_491	*RET*^{S795}	138	2.2%
JAK1_255	*JAK1*^{L511}	121	1.1%
PSMB5_91	***PSMB5*^{M104}**	**66**	**1.1%**

d

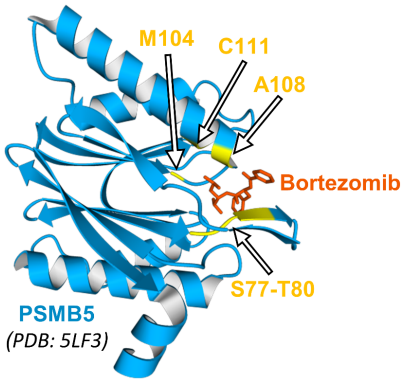


e

Mutations present in the bortezomib resistant cell pool

PSMB5 Mutation	Type	Associated sgRNA	Frequency
✓ A108T	SNV	PSMB5_94	36.2%
Intron1-del8	Deletion	PSMB5_117	34.0%
✓ A79G, T80A	MNV	PSMB5_78; PSMB5_85	9.9%
R78_A79insS*LQLT	Insertion	PSMB5_78; PSMB5_85	9.0%
A108fs	Frameshift	PSMB5_94	2.3%
✓ C111F	SNV	PSMB5_94	2.0%
✓ M104I	SNV	PSMB5_91	0.9%
T80del	Deletion	PSMB5_78; PSMB5_85	0.6%
S77A	SNV	PSMB5_78; PSMB5_85	0.6%

f



g

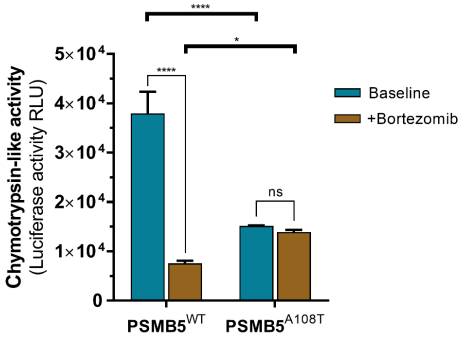


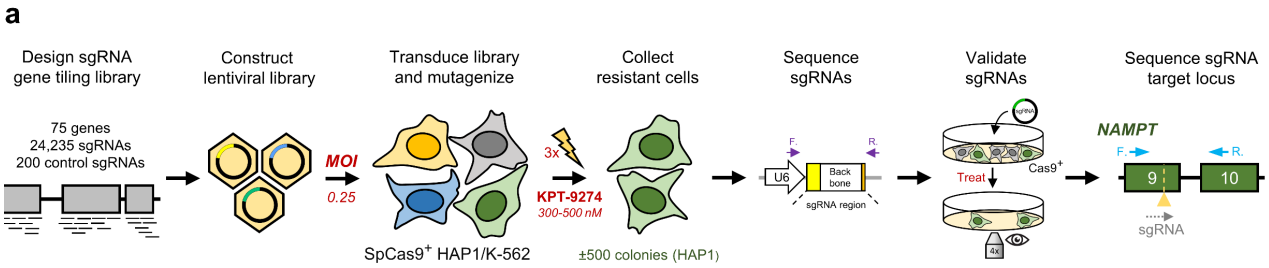
Figure 3: *validation of CRISPR-Cas mutagenesis scanning as a target deconvolution approach.*
a. Overview of the workflow for the CRISPR-SpCas9-based target identification screen used for bortezomib. Two lentiviral sgRNA sublibraries A and B, together tiling 115 genes targeted by FDA-approved anticancer drugs, were constructed (A: *ABL1* through *HDAC9*; B: *IFNAR1* through *VEGFB*). Sublibrary B was transduced into HAP1 cells stably expressing SpCas9. Transduced cells were then enriched for with puromycin and treated with 30 nM bortezomib for a period of 19 days. sgRNAs in the surviving cells were then sequenced using next generation sequencing and validated independently.
b. Cell viability of parental and the pool of bortezomib-resistant HAP1 cells in the presence of different concentrations of bortezomib. Data points represent mean ± s.d. obtained from three experiments performed in triplicate.
c. Enriched sgRNAs in cells after treatment with 30 nM bortezomib as determined by EdgeR analysis of next-generation sequencing data. Each dot represents a different sgRNA. A value of 1 was added to each read count to facilitate $\log_2$ transformation. An overview of the sgRNA hits enriched in the bortezomib-surviving cells is shown in the table (FC: fold change).
d. Validation of the sgRNA hits by assessing their ability to induce bortezomib resistance. Individual sgRNAs were cloned into an expression vector and transfected separately into SpCas9$^+$ HAP1 cells. Cells were treated with 20 nM bortezomib for 8 days and then counted using trypan blue exclusion. An sgRNA targeting the *AAVS1* safe harbor locus was used as negative control. The columns represent mean ± s.d. obtained from two separate experiments.
e. PSMB5 amino-acid variants present in the pool of bortezomib-resistant cells as determined by CrispRVariants analysis of targeted amplicon sequencing. Amino-acid mutations that are known to confer drug resistance to bortezomib are marked with a blue checkbox.
f. Cocrystal structure of PSMB5 with bortezomib. The residues found mutated in the pool of resistant cells map to the bortezomib binding site and are highlighted in yellow. The crystal structure was obtained from the protein data bank (PDB: 5LF3)[437].
g. PSMB5 chymotrypsin-like activity as measured in parental HAP1 and a single-cell-derived PSMB5^{A108T} mutant obtained from the pool of resistant cells. Cells were untreated or treated for 2hr with 12.5 nM bortezomib. Values represent mean ± s.d. obtained from a single experiment performed in triplicate; *p-value < 0.05, ****p-value < 0.0001 (two-way ANOVA with Bonferroni correction for multiple testing), ns: not significant, RLU: relative light unit.

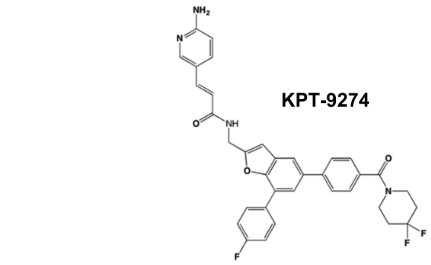
In more detail, the most enriched mutation (A108T) alters a residue known to bind bortezomib directly. Other well-known bortezomib drug resistance mutations (C111F and M104I) were identified, as well as new and earlier described mutations around threonine 80 (S77A, A79G, T80A, T80del), which is also implicated in the binding of bortezomib. One peculiar intronic mutation (deletion of the nucleotides 5'-GCCCCCTT) was identified 11bp downstream of the exon1-intron1 boundary. Although we could not validate the associated sgRNA (PSMB5_117) (**Fig. 3d**), another recent study using targeted mutagenesis of *PSMB5* also identified a mutation in the first intron of *PSMB5* in bortezomib resistant cells[303]. To further investigate the effect of bortezomib on mutated cells, we obtained a single-cell derived A108T mutant from the pool of resistant cells and examined chymotrypsin-like proteasome activity in these cells (**Fig. 3g**). Baseline proteasome activity in mutant cells was decreased when compared to the wild-type ($p<0.0001$) but treatment with bortezomib did not inhibit proteasome activity, while it was severely impaired in the parental wild-type ($p<0.0001$). Together, these results clearly validate this tiling library and demonstrate the feasibility of a SpCas9-directed mutagenesis scanning strategy for target identification on a larger scale.

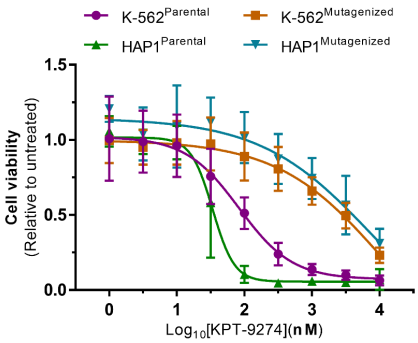
Identification of the cellular target of KPT-9274

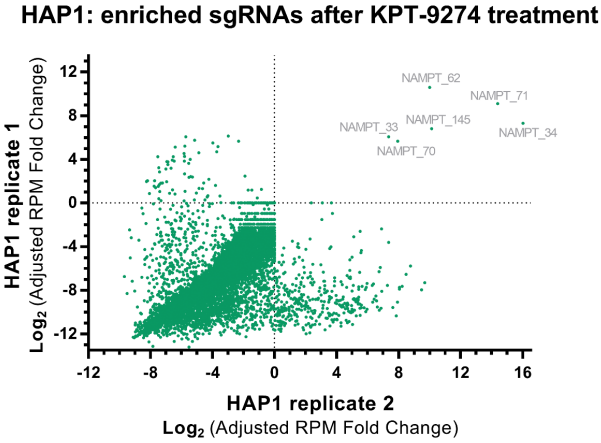
Next, we applied the approach to the clinical stage anticancer compound KPT-9274, an orally bioavailable small molecule with potent activity against different cancer types that has recently entered phase 1 clinical studies. Based on chemical proteomics experiments, KPT-9274 has been suggested to interact with the p21-associated kinase 4 (PAK4)[438] and based on experimental data, KPT-9274 has also been suggested to inhibit nicotinamide phosphoribosyl transferase (NAMPT)-mediated NAD+ biosynthesis inside the cell[439]. However, the causal association between these activities and cancer cell sensitivity to KPT-9274 remains to be investigated. Therefore, we applied our sgRNA tiling library containing 75 target genes of investigational antineoplastic drugs, which includes guides against *PAK4* and *NAMPT*, to HAP1 and chronic myelogenous leukemia K-562 cells stably expressing SpCas9 (**Fig. 4a** and **Supplementary Data 4**). Cells were pulsed three times with KPT-9274 (**Fig. 4b**) and colonies that were resistant to KPT-9274 (**Fig. 4c**) appeared within 7-14 days post treatment. In both cell lines, six sgRNAs, all targeting *NAMPT*, were enriched in the resistant cell pools (**Fig. 4d** and **e**, and **Supplementary Data 6**). From these, five were identical in both cell lines and some of these guides had overlapping target sequences (**Fig. 4f**). One sgRNA, targeting *NAMPT* codon G383, was by far the most enriched in both HAP1 and K-562 resistant cell pools and represented between 22% and 38% of total sgRNAs. In addition, an sgRNA targeting *NAMPT* that was present as little as 0.02% could be detected. When transfected separately, all enriched sgRNAs could be confirmed to confer resistance to KPT-9274 (**Fig. 4g**).

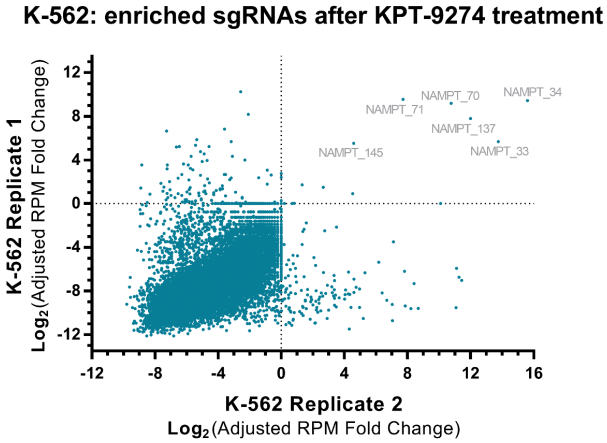
The same KPT-9274 resistance could also be elicited in pancreas carcinoma MIA PaCa2 and multiple myeloma OPM-2 cells by cotransfection of the respective NAMPT sgRNAs and SpCas9 (**Supplementary Fig. 9a**). The original KPT-9274 resistant cell pools were also cross-resistant to the prototype NAMPT inhibitor FK866[440] (**Supplementary Fig. 9b** and **c**), and the *NAMPT* targeting sgRNAs identified in our screen were able to confer HAP1 cells with resistance to FK866 (**Supplementary Fig. 9d**). Moreover, addition of β-nicotinamide mononucleotide, the product of the enzymatic reaction catalyzed by NAMPT, to parental wild-type cells provided similar protection against KPT-9274-induced cytotoxicity as the *NAMPT* targeting sgRNAs from our screen (compare **Fig. 4c** with **Supplementary Fig. 9e**), further pinpointing NAMPT is a key cellular target of KPT-9274. We then identified the underlying resistance mutations present in the mutagenized population of cells obtained from the screen. As some of the enriched sgRNAs had overlapping target sequences (NAMPT_34 and NAMPT_33; NAMPT_137 and NAMPT_145), only four domains within the NAMPT protein were effectively targeted. Deep sequencing of *NAMPT* at these four domains revealed a variety of mutations (**Fig. 4h** and **Supplementary Data 7**). The G383del mutation was by far the most prevalent, ranging from 10% to 52%, and was the only mutation identified throughout all samples. The P236_V237del mutation was only present in the two replicates from the HAP1 experiment, while the G381_G384del mutation was only present in the K-562 replicates. Sequencing of the first *NAMPT* exon, which is targeted by the NAMPT_137 and NAMPT_145 sgRNAs, proved to be difficult due to the high GC content in the flanking introns. In some samples we did manage to identify the I11_K19del, S17del and Y18del mutations, although we could not quantify these mutations effectively. Other in-frame mutations, including P238_Y240insAAEHS, S275T, S382_G385del, and some frameshift mutations, were identified in individual replicates. Frameshift mutations are most probably present in hemizygous cells since knock-out of *NAMPT* is lethal in HAP1 and K-562 cells (**Supplementary Fig. 10d**). We also identified some low prevalent NAMPT mutations in individual replicates that did not localize to the target sites of the enriched sgRNAs. Three of these mutations (G217E, G217V, G217R) alter the G217 residue that has been identified earlier to confer resistance to NAMPT inhibitors[441]. These mutations were found in one of the two replicates and were likely already present in the heterogenous parental population or arose during the experiment. Finally, one known SNP (S301S), rs2302559 (dbSNP), was also detected in both replicate HAP1 screens.





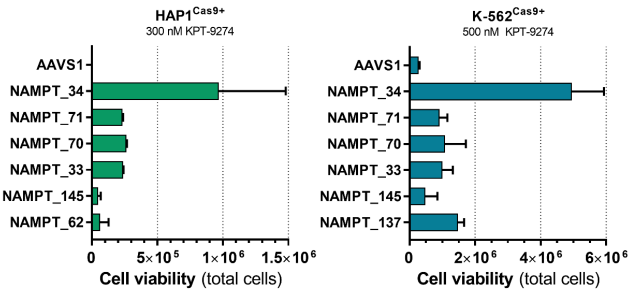






f

		HAP1 (averages)		*K-562 (averages)*	
sgRNA ID	**Cut Site**	**Log_2(FC)**	**Frequency**	**Log_2(FC)**	**Frequency**
NAMPT_34	*NAMPT^{G383}*	11.66	35.65%	12.53	24.60%
NAMPT_71	*NAMPT^{P238}*	11.75	2.86%	8.63	0.06%
NAMPT_70	*NAMPT^{Y240}*	6.81	0.48%	9.98	1.08%
NAMPT_33	*NAMPT^{S382}*	6.70	0.02%	9.71	0.70%
NAMPT_145	*NAMPT^{Y18}*	8.46	6.72%	5.06	1.18%
NAMPT_62	*NAMPT^{S275}*	10.29	1.20%	-2.36	0.00%
NAMPT_137	*NAMPT^{A14}*	-4.04	0.00%	9.89	16.25%



h

Most prevalent NAMPT mutations detected in HAP1 replicate #1

NAMPT mutation	Type	Associated sgRNAs	Frequency
G383del	Deletion	NAMPT_33; NAMPT_34	51.90%
P238_Y240insAAEHS	Insertion	NAMPT_70; NAMPT_71	2.53%
G217V	SNV	-	1.91%
P236_V237del	Deletion	NAMPT_70; NAMPT_71	1.64%
P236fs	Frameshift	NAMPT_70; NAMPT_71	1.02%
G239V	SNV	NAMPT_70; NAMPT_71	0.26%
G239_Y240insY	Insertion	NAMPT_70; NAMPT_71	0.24%
S275fs	Frameshift	NAMPT_62	0.17%
G239D	SNV	NAMPT_70; NAMPT_71	0.15%
S301S	Silent SNV	-	100.00%

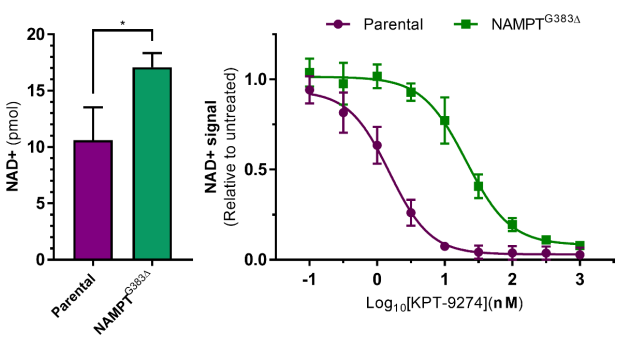


Figure 4: *CRISPR-Cas9 mutagenesis scanning reveals NAMPT as the cellular target of KPT-9274.*
a. Experimental workflow for the CRISPR-SpCas9-based target identification screen for KPT-9274. HAP1 or K-562 SpCas9$^+$ cells were transduced with the investigational target gene library, enriched by puromycin selection and treated with KPT-9274. Surviving cells were harvested and the sgRNAs sequenced. The screens were repeated once for each cell line.
b. Chemical structure of KPT-9274.
c. Resistance profile of parental cells and transduced cell populations that survived treatment with KPT-9274. Cell viability was measured in the presence of increasing concentrations of KPT-9274 and was adjusted to the untreated control. Data points represent mean ± s.d. obtained from two (HAP1) or five (K-562) experiments performed in triplicate.
d. Enriched sgRNAs present in transduced HAP1 cells after treatment with KPT-9274 (300 nM). Fold change was determined by taking the $\log_2$ of the reads per million (RPM) after treatment added by 1 (adjusted) and then divided by the adjusted RPM before treatment. Read counts were determined with EdgeR and each dot represents an sgRNA.
e. Enriched sgRNAs present in transduced K-562 cells after treatment with KPT-9274.
f. Overview of enriched sgRNAs present in transduced cells after treatment with KPT-9274. Values represent averages of replicate experiments for each cell line. FC: adjusted fold change.
g. sgRNAs identified in the screen were validated by assessing their ability to induce drug resistance. Each sgRNA was transfected separately and cells were treated with KPT-9274 for four (HAP1) or six (K-562) days. Surviving cells were counted using trypan blue exclusion. A safe-targeting sgRNA (*AAVS1*) was included as negative control. The columns indicate mean ± s.d. obtained from two experiments.
h. Prevalent NAMPT mutations detected in KPT-9274-resistant cells from HAP1 replicate #1.
i. NAD+ levels in parental and HDR-edited NAMPT^{G383del} HAP1 cells treated with increasing concentrations of KPT-9274. The left panel shows untreated baseline NAD+ levels. Two experiments were performed in duplicate and values represent mean ± s.d. *p-value < 0.05 (two-tailed Mann–Whitney U test).

To validate our findings, we reinstalled the G383del mutation, which was the only mutation detected in all HAP1 and K-562 replicates, into its native locus by CRISPR-induced HDR of parental HAP1 cells (**Supplementary Fig. 10a**). HDR-edited cells were resistant to KPT-9274 treatment and were also cross-resistant to FK866 (**Supplementary Fig. 10b** and **c**), again providing further evidence for NAMPT as the key cellular target of KPT-9274. To examine the function of NAMPT protein containing the G383del mutation, cellular NAD+ levels were measured in absence and presence of KPT-9274. Interestingly, baseline NAD+ levels were slightly elevated in mutant cells when compared to parental cells (**Fig. 4i** left panel) and treatment of parental cells with KPT-9274 decreased NAD+ levels rapidly at low nanomolar concentrations, while the G383del mutation conferred resistance (**Fig. 4i** right panel). To unambiguously validate our results and corroborate the validity of the conclusions taken from our mutagenesis scanning approach, we co-crystallized a NAMPT dimer in complex with KPT-9274 (**Fig. 5a** and **Supplementary Fig. 11**) (PDB: 5NSD) and explained the identified resistance mutations by modeling them into the structure (**Fig. 5b**). Crystallographic analysis of the binding mode revealed that two KPT-9274 molecules are bound to the NAMPT dimer (**Fig. 5a**) and their binding mode is largely similar to the one of the previously reported prototype NAMPT inhibitor FK866 (**Fig. 5b** left lower panel)[440].

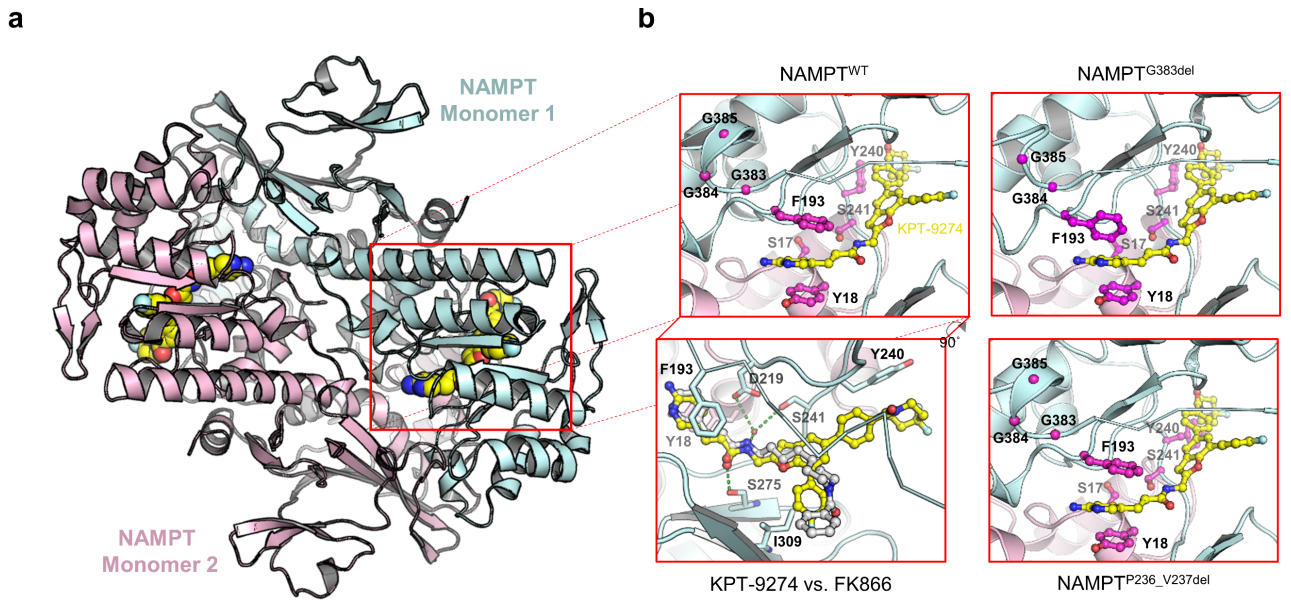


Figure 5: *co-crystal structure of NAMPT dimer with KPT-9274.*
a. The crystal structure of NAMPT dimer with KPT-9274 is depicted as a cartoon with the NAMPT monomers shown in light blue and pink and KPT-9274 as yellow spheres.
b. A sideways blow-up (top left) shows the drug-binding site with KPT-9274 depicted in yellow carbons. A second blow-up (down left) shows the overlap between KPT-9274 and the NAMPT inhibitor FK866. The pi-stacking of KPT-9274 between Y18 and F193 matches the binding mode of FK866 and important hydrogen bonds, depicted as green dashed lines (between S275 and protein-bound water coordinated by D219 and S241), overlap between the two compounds. The right panels show the influence of the G383del (top) and P236_V237del (bottom) resistance mutations as modelled with the Molecular Operating Environment (MOE).

In more detail, the amino-pyridinyl of KPT-9274 is coordinated via pi-stacking interactions between tyrosine 18 and phenylalanine 193. Furthermore, the amide linker is well coordinated via hydrogen bonding interactions and the free electron pairs of the amide oxygen are stabilized by a hydrogen bond originating from serine 275. The nitrogen forms a hydrogen bond with a water molecule that is interacting with the carboxyl group of aspartate 219 and the hydroxyl group of serine 241. At the solvent exposed regions of the binding site, the interactions between KPT-9274 and NAMPT differ from FK866 as the KPT-9274 core consists of a rigid benzofuran moiety, contrary to the flexible carbon chain in FK866 (compare **Fig. 4b** and **Supplementary Fig. 9b**). KPT-9274 binds two solvent exposed patches whereas FK866 only occupies a single hydrophobic site. The overlapping hydrophobic patch outlined by hydrophobic sidechains of isoleucine 309, isoleucine 351, and alanine 379 is occupied by the aromatic fluorophenyl group of KPT-9274. The second solvent exposed binding site, which is only covered by KPT-9274, consists of tyrosine 188 and tyrosine 240 that interact with the difluoro-1-piperidinyl carbonyl phenyl group via Van der Waals interactions and a hydrogen bond between the carbonyl group and the amino group of lysine 216. Importantly, the sgRNAs identified in our screen target close to residues (tyrosine 18, tyrosine 240, serine 241, serine 275 and alanine 379) that are confirmed by the co-crystal to take part in the binding of KPT-9274, demonstrating the power of the CRISPR-Cas9 mutagenesis scanning approach. Finally, to investigate the influence of the discovered resistance mutations on the NAMPT structure and ligand binding, the G383del and P236_V237del mutations were modelled (**Fig. 5b** right panels). Deletion of G383 causes a conformational change of the G383-385 stretch, which destabilizes the conformation of F193 and thereby prevents pi-stacking with KPT-9274. The P236_V237del mutation causes a displacement of the loop containing the Y240 residue, resulting in steric incompatibility with KPT-9274.

AsCpf1 extends the resolution of the methodology

The potential of the SpCas9 mutagenesis scanning approach may be limited by the availability of NGG PAMs at or nearby the resistance hot spots of the investigated drug. To determine if this could present as a problem, we estimated the mutational range across the coding sequence of the target genes in our SpCas9 libraries using a computational analysis. For this purpose, we considered an effective mutational range of 9bp around the SpCas9 cut site (**Fig. 6a**). This range was based on the observed nucleotide variant occurrence at sgRNA target sites determined earlier (**Fig. 1e** and **g** and **Supplementary Figs. 1-4**) and on results obtained by analyzing the outcomes of NHEJ repair of SpCas9-mediated cutting in human cells of *KIF11*, *XPO1* and the *AAVS1* safe harbor locus without drug selection (data not shown). Applying this range to the target genes in our SpCas9 tiling libraries, we estimated that 73.4 ± 11.4% (mean ± s.d.) of the coding sequences falls within the 9 bp mutational range for SpCas9, although the mutational coverage varies greatly between genes (**Fig. 6b** and **c**). To mitigate this restraint, we calculated the mutational coverage of the target genes in our tiling libraries using a similar 9 bp window around the cut site for the class 2 type V CRISPR-Cas AsCpf1 (Cas12) endonuclease, which recognizes the TTTN PAM sequence[226]. Although generally the mutational coverage of individual genes by AsCpf1 was much lower than for SpCas9 (**Fig. 6c** and **d**), combination of sgRNAs targeting both SpCas9 NGG and AsCpf1 TTTN PAMs expanded the overall mutational coverage, highlighting some parts of the coding sequence that are not covered by SpCas9 can be mutated using AsCpf1 (**Fig. 6c** and **d**).

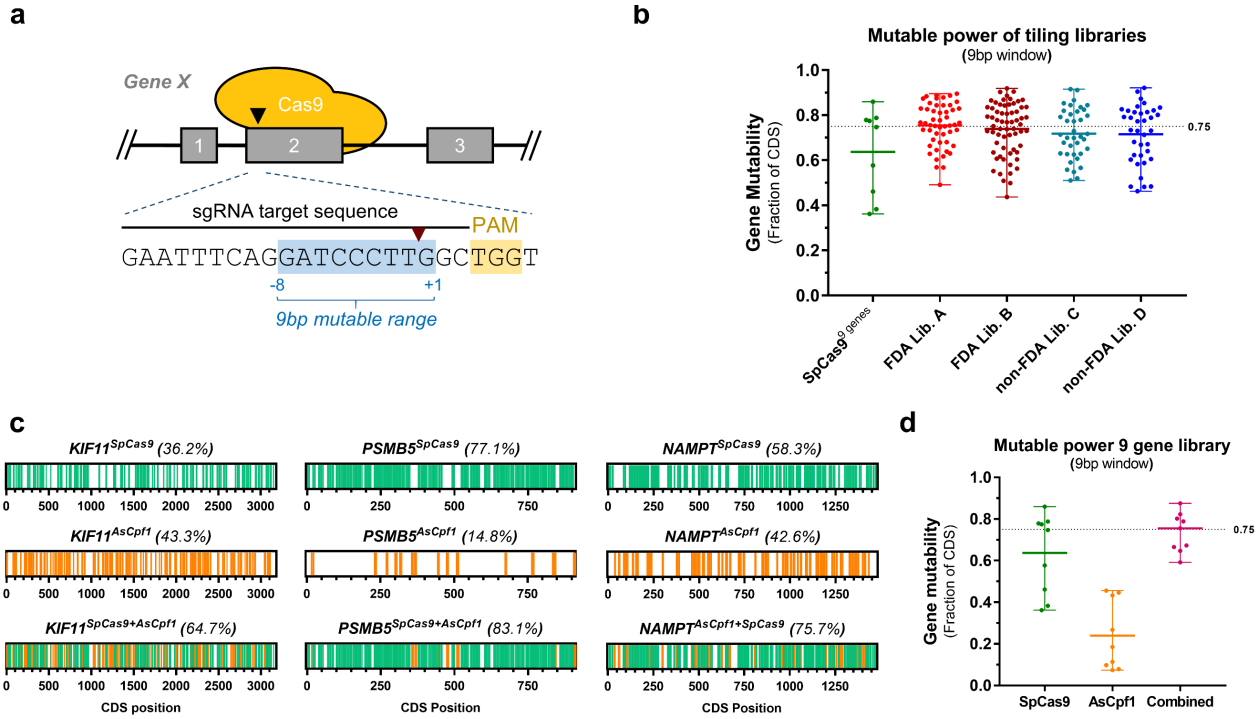
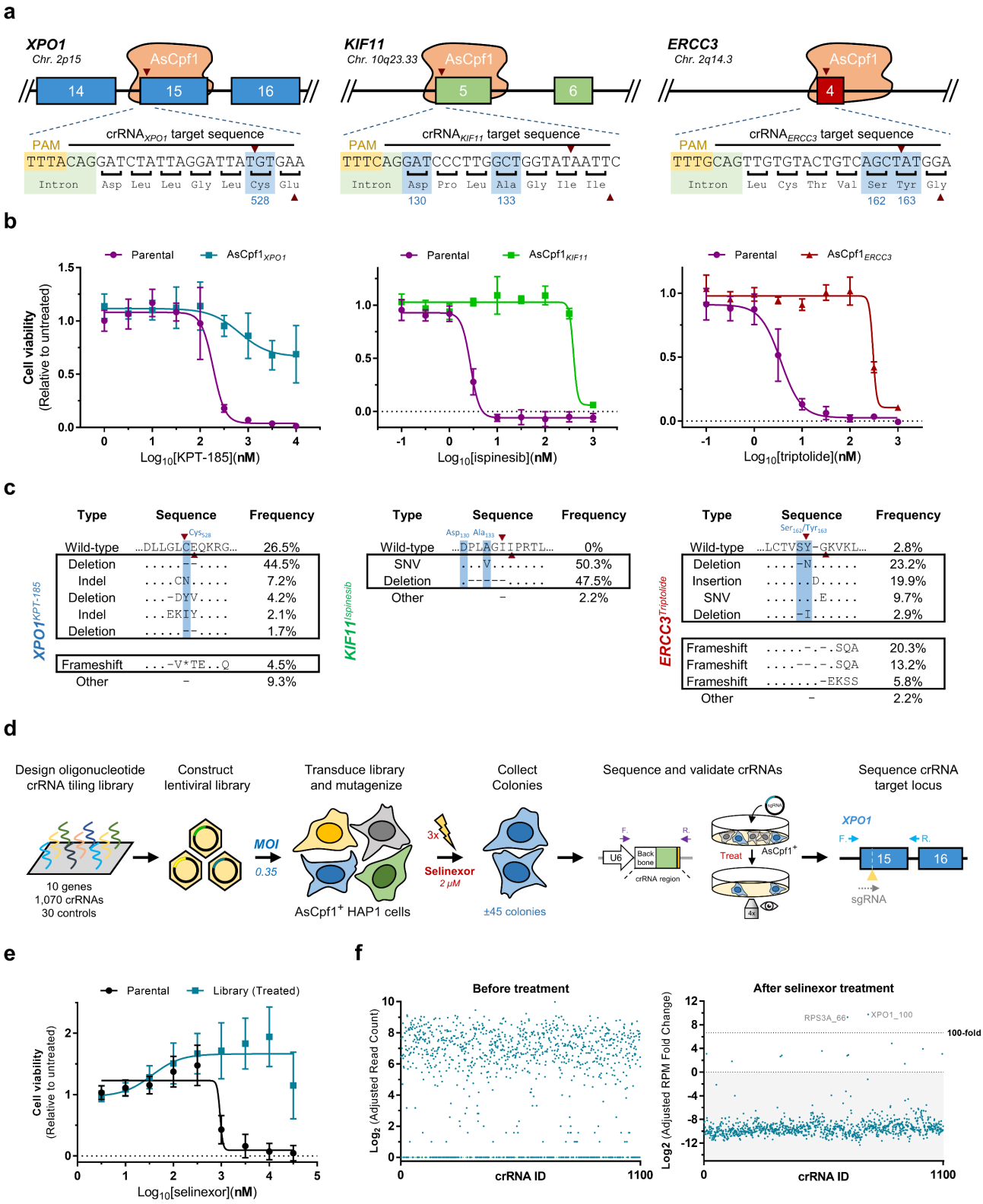


Figure 6: *mutational range across the coding sequence of target genes in the tiling libraries.*
a. Illustration of the estimated mutable range for a given sgRNA. Red arrowhead shows SpCas9 cut site.
b. Scatter dot plot showing the calculated mutational coverage for the coding sequences of each target gene in our SpCas9 tiling libraries. Values were calculated using the 9bp mutable range. Each dot represents a different gene. The bold horizontal line indicates the mean gene mutability for a given library, while the whiskers illustrate the total range between lowest and highest values.
c. Example of the calculated mutational coverage for SpCas9 (green) and AsCpf1 (orange) of three target genes (*KIF11*, *PSMB5* and *NAMPT)*. Colored regions indicate the region of the coding sequence that falls within a mutable range, while white areas indicate the region falls outside of the mutable range.
d. Scatter dot plot showing the calculated mutational coverage for the 9 gene library tiled with guides for SpCas9 (green), AsCpf1 (orange) or SpCas9 and AsCpf1 combined (purple). A clear increase of the average mutational coverage is visible when combining guide RNAs for both SpCas9 and AsCpf1.



g

Enriched crRNAs after selinexor treatment

crRNA ID	Cut site	Fold Change	Frequency
XPO1_100	*XPO1*^{C528}	875	71.9%
RPS3A_66	*RPS3A*^{F138}	624	19.9%

h

XPO1 mutations detected at the crRNA cut site

XPO1 mutation	Type	Associated crRNA	Frequency
C528_E529del	Deletion	*XPO1_100*	43,96%
G526E,L527L,C528I,E529L	MNV	*XPO1_100*	6,32%

Figure 7: *the AsCpf1 endonuclease facilitates rapid selection of functional resistance mutations.*
a. An overview of AsCpf1 crRNAs targeting *XPO1*, *KIF11*, and *ERCC3* at their resistance hotspot residues. The AsCpf1 cleavage site generates an overhang and is denoted by red arrowheads.
b. Cell viability of wild-type and polyclonal mutagenized resistant cells in the presence of increasing concentration of drug. Values are shown relative to untreated controls. Data points represent mean ± s.d. obtained from two experiments performed in triplicate.
c. Amino-acid sequence variants determined by targeted amplicon-sequencing analysis of the resistant cells with CrispRVariants. Only variants with a frequency ≥1% are shown.
d. An overview of the CRISPR-AsCpf1 tiling crRNA library approach used for target identification of selinexor. A lentiviral tilling library targeting 10 different genes was constructed and transduced in cells stably expressing AsCpf1, which were first selected with puromycin and subsequently treated with selinexor (2 μM). crRNAs present in resistant colonies were identified by next-generation sequencing after which the genomic locus targeted by these crRNAs was sequenced to identify the resistance conferring mutations.
e. Cell viability in the presence of different concentrations of selinexor of wild-type and resistant cells obtained after library transduction and selinexor treatment. Data was obtained by performing three experiments in triplicate and data points represent mean ± s.d.
f. Representation of the different crRNAs present in the cells before (after puromycin selection) and after treatment with selinexor as determined by EdgeR analysis of next-generation sequencing data. A value of 1 was added to each read count to facilitate $\log_2$ transformation. Each dot represents a different crRNA.
g. crRNA hits enriched over 100-fold in the selinexor-resistant cell pool after AsCpf1-mediated mutagenesis screen. Read counts were determined using EdgeR.
h. Amino-acid variants detected at the XPO1 C528 locus of resistant cells. Alterations were uncovered by CrispRVariants analysis of next-generation sequencing data.

To investigate whether the approach is compatible with AsCpf1, we selected AsCpf1 crRNAs targeting TTTN motifs around the same codons for *KIF11*, *ERCC3* and *XPO1* as described in the first figure for SpCas9 (**Fig. 7a** and **Supplementary Fig. 12a**). These were transfected along with AsCpf1 in HAP1 cells that were subsequently treated with the respective compounds. Colonies that formed within a few days of treatment were drug resistant to the specific treatment (**Fig. 7b** and **Supplementary Fig. 12b**) and contained mutations at the known resistance hot spots (**Fig. 7c** and **Supplementary Fig. 12c**). We next designed a lentiviral crRNA tiling library, similar to the 9 gene SpCas9 tiling library (**Supplementary Fig. 8a**), spanning all possible TTTN PAM sites in the exons from 10 genes (1,100 crRNAs, **Supplementary Data 8**) and applied it to selinexor (**Fig. 7d**). Colonies rapidly appeared and were resistant to selinexor (**Fig. 7e**). One crRNA targeting codon C528 in *XPO1*, the anchor point of selinexor, was highly enriched (71.9%) in the resistant cell pool (**Fig. 7f**, **g**, and **Supplementary Data 9**). Next generation sequencing of the targeted *XPO1* locus revealed protein variants that mainly consisted of a deletion of residues C528 and E529 (**Fig. 7h**), identical to what we observed for the single crRNA (**Fig. 7c**). One other crRNA targeting the 40S Ribosomal Protein S3a (RPS3A) was also enriched but could not be validated when transfected individually (data not shown). In addition, single-cell derived clones from the original resistant cell pool revealed that this *RPS3A* targeting crRNA always co-appeared with the $XPO1^{C528}$ crRNA (**Supplementary Fig. 13a** and **b**), suggesting these cells were transduced with two lentiviral particles and that decreasing the MOI will reduce the false positive rate. To conclude, these results demonstrate that the CRISPR-based mutagenesis scanning approach is compatible with the AsCpf1 endonuclease, showing the approach can be tailored to other CRISPR-Cas endonucleases to increase the resolution of the method.

Discussion

One of the major challenges in drug discovery is the identification of the cellular target of a valuable bioactive small molecule. Currently, a wide range of target deconvolution technologies is available[38, 39, 44], but the gold standard for target confirmation is the identification of mutations that confer drug resistance. Knowledge of resistance mutations is not only used as proof for target engagement but can also provide functional insight in the molecular biology of the target protein and can be used to rescue phenotypes from chemical genetic screens[369]. Here, we show that genetic variation introduced by targeted CRISPR-Cas-induced NHEJ facilitates rapid selection of in-frame drug resistance mutations in cancer cells. Because this approach exploits an error-prone DNA break repair mechanism, it generates a wide diversity of heterogeneous mutations. We demonstrate this simple and fast method can be applied as a genetic screening strategy using large tiling libraries to identify the direct cellular target protein of a small molecule. An important strength of this targeted mutagenesis scanning method is that the single guide RNA sequences directly annotate the genomic sequence that contains the drug resistance-conferring mutations, avoiding the need for whole exome sequencing and complex deconvolution endeavors to uncover the relevant resistance mutations. In addition, a replicate screen allowed for the detection of extremely low-frequency (0.02%) sgRNA hits with high confidence (**Fig. 4d** and **e**) allowing for the identification of infrequent mutational hotspots conferring drug resistance. Nevertheless, a single screen is sufficient to detect important drug resistance conferring sgRNAs with associated mutations as demonstrated with bortezomib (**Fig. 3**). However, some false positive sgRNAs were detected in this single screen. This can be explained by co-transduction of the resistance conferring sgRNA with another sgRNA in a small fraction of the cells, happening even at low multiplicity of infection.

Another source of noise can arise from unrelated sgRNAs that were transduced in a very small fraction of cells that already contain resistance conferring mutations due to the heterogeneity and genetic instability of the cancer cell line. These false positive hits can be eliminated by a replicate screen or by a simple transfection experiment of the individual sgRNAs (**Fig. 3e**). Another potential pitfall arises from the possibility of discovering activating mutations in other proteins besides the target protein, as these may obscure the identification of the direct target itself. However, although in our screens we did not observe such activating mutations, it remains to be seen how often they will occur and how they will affect the methodology. Of note, resistant cells acquired by the CRISPR-Cas mutagenesis approach were commonly hemizygous for the resistance mutation (**Supplementary Fig. 1b**). Therefore, the approach allows for the discovery of recessive mutations avoiding the need for haploid cells. This is further illustrated by triptolide resistance, for which a wild-type allele is sufficient for sensitivity[121], in multiploid HCT 116 and HL-60 cells (**Supplementary Fig. 6** and **7**).

Recently, the targeting of cytidine deaminases with nuclease-deficient Cas9 has been shown to induce site-specific mutagenesis without introducing a DSB to obtain drug resistance[303, 304]. This experimental set up, a form of base editing, is different from the here described NHEJ-based mutagenesis approach as, innately, the base-editing approaches cannot derive deletions or insertions and excel at generating single point mutants. However, we observed in-frame deletions and insertions can provide a major mechanism for drug resistance, even when localized to functionally important protein domains. Moreover, these base-editing approaches are less probable to recover recessive drug resistance mutations as this requires the same mutation to be present in both alleles. In addition, although the currently most hyperactive variant of the base editing variants can cover a mutational hotspot region of around 100 bases, it is limited to the introduction of an average of 1.32 base substitutions per read[303]. The NHEJ-based approach described here covers a smaller hotspot region but generates larger regions of genetic variation of up to 17 bases per read, although many of these mutations include deletions. Another major difference between base editing and NHEJ-based mutagenesis methods is the ability to multiplex. Due to the nature of the NHEJ-based methodology, it will be

almost impossible to perform multiplex mutagenesis in a single cell, especially if performed on the same chromosome. Indeed, CRISPR-Cas-mediated induction of double strand breaks will lead to the formation of large deletions or even translocations when targeting multiple loci in the same cell. Base editors clearly have an advantage here, as they do not induce double strand breaks and have already been shown to facilitate multiplex mutagenesis at different sites in the same cell[303].

Given the caveat of different experimental setups between the methods, we can compare CRISPRres to one of the hyperactive base-editing tools (CRISPR-X, dCas9-AID*Δ) based on published data with bortezomib[303]. Both methods have used the bortezomib-PSMB5 drug-target pair to obtain drug resistance mutations by screening. Mutations at the A108 residue known to be involved in binding bortezomib were discovered by both approaches (CRISPRres: A108T; CRISPR-X: A108T, A108V). CRISPRres also identified well-known bortezomib mutations nearby this A108 residue (C111F, M104I), while the CRISPR-X screen did not discover these mutations. Both methods identified novel mutations near residue T80, which also binds bortezomib directly (CRISPRres: A79G, S77A; CRISPR-X: A74V, R78M, R78N, A79T, A79G, G82D), while CRISPRres also unraveled mutations of the T80 residue (T80A, T80Δ). On the other hand, the CRISPR-X screen discovered a G242D mutation distal to the bortezomib binding pocket, which was not identified by our approach. Interestingly, both methods discovered a peculiar mutation in the first intron of *PSMB5*, although both studies were unable to validate these intronic mutations. In conclusion, both studies discovered resistance mutations that were similar, although it is important to note the CRISPRres screen probed a pool of 64 genes simultaneously, while CRISPR-X was only applied on a single gene (*PSMB5*).

Using the CRISPRres approach, we also revealed NAMPT as the key cellular target of the clinical inhibitor KPT-9274 by showing that the cytotoxic activity of KPT-9274 is dependent on inhibition of NAMPT in the used cell lines. Tumor cells have been shown to be more dependent on the NAD+ salvage pathway catalyzed by intracellular NAMPT[442] and overexpression of NAMPT has been observed in different cancer types and has been implicated to confer resistance to classic chemotherapeutics, whereas inhibition sensitizes cancer cells to oxidative stress and DNA damaging agents[443]. Thus, inhibition of NAD+ biosynthesis facilitated by NAMPT shows potential as anticancer therapy. From this perspective, multiple efforts have led to the discovery of NAMPT inhibitors[168, 444]. These inhibitors induce delayed apoptosis or autophagy by gradual NAD+ or ATP depletion, loss of plasma membrane potential and possibly by lowered activity of PARPs and SIRTs[445]. Although some NAMPT inhibitors have entered clinical trials as single agent (FK866/APO866 and GMX1777/GMX1778), they failed to progress to phase 2 due to on-target dose-limiting toxicity (mainly thrombocytopenia) and lack of clinical efficacy[446]. This might partially be explained by the still poorly understood aspects of NAMPT biology in bigger biological systems, which might have led to the selection of inappropriate indications or patient populations. In addition, some of these inhibitors have been noted to have relatively poor pharmacokinetics[446]. KPT-9274 represents a new class of NAMPT inhibitors and shows promising pharmacokinetics (Karyopharm Therapeutics). Different studies have already shown KPT-9274 has promising anticancer activity *in vitro* and in preclinical models[438, 439, 447, 448], and it has entered phase 1 clinical trials (ClinicalTrials.gov: NCT02702492). These trials are designed to include an arm with co-treatment of KPT-9274 with nicotinic acid (NA) as the existence of a secondary, NAPRT1-dependent NAD+ salvage pathway, which converts NA to NAD+, may mitigate the previously observed dose-limiting toxicity due to NAMPT inhibition in normal tissue cells. Indeed, unlike many tumor types, normal tissue cells express functional NAPRT1 and can generate NAD+ from NA. Therefore, this co-treatment may be one of the keys to the success of KPT-9274 in the clinic.

Conclusion

Altogether, our findings demonstrate that the localized genetic variation generated by CRISPR-mediated NHEJ repair can be exploited to screen essential genes for gain-of-function mutations. We establish this mutagenesis scanning approach as a genetic screen to identify the molecular target of chemical compounds inhibiting essential proteins and the method is therefore complementary to loss-of-function screens. We envisage this genetic screen can either be applied on a list of candidate genes identified after a first round of target deconvolution by e.g. affinity chemical proteomics, overexpression screens, or comparative studies such as CMap[449], or it can be applied on a predefined shortlist of targets of interest. Indeed, it allows to rapidly select from a primary screen those hit molecules that target a protein or pathway of interest. Nevertheless, even when no *a priori* knowledge on a potential target of a hit molecule is available, the current format of the method allows coverage of all essential genes (±2,000 for a given cell line)[185, 268, 269] with 20-25 tiling libraries for target discovery. Finally, we have illustrated the application of this genetic screen for the identification of drug-target interactions using cellular toxicity as phenotypic selection, but it may also be applicable to other phenotypic reporter assays and even to protein engineering in a native context.

Methods

Cell culture

HAP1 cells were obtained from Horizon Discovery. HCT 116, MIA PaCa2, and K-562 cells were obtained from American Type Culture Collection (ATCC), HL-60 cells from Sigma Aldrich, and OPM-2 from DSMZ. SpCas9 and AsCpf1 expressing cells were generated in house. HAP1, HL-60, and K-562 cells were grown and passaged every 2–3 days in IMDM. HCT 116 cells were grown in McCoy's 5A medium, MIA PaCa2 cells were cultured in DMEM, and OPM-2 cells in RPMI 1640. All media were supplemented with 10% FBS and 20 μg/mL gentamicin. Cells were incubated at 37 °C and 5% CO_2 and were regularly checked for mycoplasma contamination using the Venor GeM OneStep PCR kit (Minerva Biolabs).

Compounds

KPT-185, selinexor (KPT-330), KPT-9058, and KPT-9274 were provided by Karyopharm Therapeutics (Newton, MA). Ispinesib, triptolide, bortezomib, β-nicotinamide mononucleotide, and FK866 were obtained from SelleckChem. All compounds were dissolved in DMSO, except for FK866, which was dissolved in ethanol.

DNA constructs

The plasmid expressing humanized SpCas9 was obtained from Labomics. The plasmid expressing humanized AsCpf1 was obtained from Addgene (69982). Plasmids containing sgRNAs or crRNAs were cloned in house. The pLCKO vector used for generation of the lentiviral library was obtained from Addgene (73311). The plasmid encoding for the XPO1 reporter cargo, NLS^{SV40}-AcGFP-NES^{PKI}, was cloned in house. Single-stranded DNA oligonucleotides for use with HDR were obtained from Integrated DNA Technologies. See **Supplementary Data 10** for the sequences of the sgRNAs, crRNAs, and HDR templates.

Generation of stable SpCas9 and AsCpf1 HAP1 cell lines

Knock-in HAP1 cell lines stably expressing SpCas9 or AsCpf1 were generated using the CRISPaint principle[450]. Briefly, cells were electroporated with the Neon Electroporation System (Thermo Fisher Scientific) using 1,450V and 3 pulses of 10 ms in 10 μL Buffer R (Neon Electroporation Kit). Cells were transfected with a plasmid encoding an sgRNA targeting the C-terminal encoding region of *SDHA* (250 ng), a plasmid encoding SpCas9 and an sgRNA targeting the donor plasmid (250 ng), and a repair donor plasmid containing a PAM and sgRNA targeting site and the sequence for P2A-mCherry-T2A-SpCas9-P2A-HygroR or T2A-AsCpf1-P2A-HyrgoR

(250 ng) to stably integrate SpCas9 or AsCpf1 downstream of the SDHA housekeeping gene. K-562 cells were transfected with a plasmid encoding an sgRNA targeting the *AAVS1* safe harbor locus, a plasmid encoding SpCas9 and an sgRNA targeting the donor plasmid and a repair donor plasmid containing a PAM and sgRNA targeting site and the sequence for CMV-mCherry-T2A-SpCas9-P2A-HygroR to stably integrate SpCas9 in the *AAVS1* safe harbor locus. All cells were plated in a 6-well plate after transfection and 2 days later, cells were selected with 300 µg/mL hygromycin B for a period of 10 days and then checked for the expression of red fluorescent mCherry. SpCas9 or AsCpf1 endonuclease activity was assessed in the polyclonal mixtures by indel detection using TIDE[451] in *XPO1* after sgRNA/crRNA transfection in the respective cell line.

Transfection

Cells were transfected with the Neon Electroporation System after resuspension in Buffer R. DNA plasmids expressing the SpCas9 or AsCpf1 endonuclease and guiding RNA were added at a concentration of 37.5 ng/µL per plasmid and electroporated at 1,400–1,475 V with 3 pulses of 10 ms. Two to three days after transfection, cells were treated with the respective drug to select for resistance over 7–14 days. Surviving cells were counted using trypan blue exclusion staining or imaged with an IncuCyte© ZOOM (Essen Bioscience). For HDR, a 123–134-bases-long ssDNA oligonucleotide (850 ng) was added to the electroporation mixture in addition to the respective sgRNA and SpCas9 expressing plasmids. Two days after electroporation, cells were treated with the respective drugs for 5 days before imaging with an IncuCyte© ZOOM (Essen Bioscience). Following imaging, the HDR-template transfected cells were grown under drug selection for an additional week before any further experiments were performed.

Cell viability assays

Cell viability assays were performed by plating 3,000 HAP1, 5,000 HCT 116, K-562, or HL-60 cells in 96-well plates containing a dilution of the test compound. Cells were incubated for 72hr at 37 °C and 5% CO_2. Cell viability was then assessed with the CellTiter 96® AQueous Non-Radioactive Cell Proliferation Assay (Promega) and colorimetric signals were measured with a Safire2™ (TECAN). Assays were performed in triplicate and each experiment was repeated at least once. Obtained values were adjusted with the background signal and divided by the untreated control. Relative data values were then visualized and analyzed using a log-based 4 parameter model (GraphPad Prism). For single-guide drug resistance validation assays, 125,000 MIA PaCa2, OPM-2, or HAP1/SpCas9$^+$ and K-562/pCas9$^+$ cells were transfected with plasmids expressing individual sgRNAs, and if needed SpCas9, using the Neon Electroporation system as described above and plated into 6-well plates. Cells were treated 2–3 days after transfection with the respective compound for a period of 5–7 days, the medium was regularly refreshed, and dead cells were washed away before imaging confluency using a live cell analysis system (Essen Bioscience, IncuCyte ZOOM®).

DNA extraction and sequencing

Genomic DNA was isolated from 1 million cells with the QIAamp DNA mini kit using RNase A. For Sanger sequencing, the region of interest was amplified by PCR with the CloneAmp HiFi PCR premix (Clontech). The amplified DNA was then purified (QIAquick PCR purification kit (Qiagen)) and sequenced (Macrogen). For targeted amplicon next-generation sequencing of single-guide transfected cells, the region of interest (*KIF11*^{A133}, *XPO1*^{C528}, *ERCC3*^{D54}, *ERCC3*^{S162}, *NAMPT*^{Y18}, *NAMPT*^{S240}, or *NAMPT*^{G383}) was first amplified over 24 cycles in 25 µL PCR reactions containing 50 ng genomic DNA with the Phusion® High-Fidelity PCR Master Mix with HF Buffer (NEB) and with custom primers containing adapter regions for Nextera indexes (IDT). Amplified DNA was purified with the QiaQuick PCR purification kit (Qiagen) and 1.5–2 µL of this DNA was PCR amplified over 25 cycles with CloneAmp HiFi PCR Premix (Clontech) using

indexing primers containing P5 and P7 Illumina adapters in 25 µL reactions to index the samples. Indexed samples were purified using magnetic Agencourt AMPure XP beads (Beckman Coulter) and eluted in TE buffer. Samples were then diluted to 2–4 nM and pooled to form the initial library. This library was then denatured and diluted according to the instructions for paired-end sequencing on a MiSeq (Illumina) with a MiSeq V2–300 or 500 cycles kit (Illumina) and 10% PhiX v3 (Illumina) spike-in. For a list of primers see Supplementary Data 10.

Analysis of next-generation sequencing data

FastQ files obtained after MiSeq sequencing were demultiplexed with the MiSeq Reporter software (Illumina). Demultiplexed and paired reads were trimmed, filtered, and then aligned to the reference amplicon in Geneious (v9, Biomatters). To obtain haplotypes present in drug-resistant samples, bam files were analyzed with the CrispRVariants[452] package run in RStudio by defining a 35–140 bp spanning region across the endonuclease cut site, as defined by pre-analysis of localized variants within Geneious.

For this purpose, the CrispRVariants "readtotarget" input was run with parameters "upstream.snv" (15–70) and "downstream.snv" (15–70) on corresponding paired end sequencing reads to allow for haplotype determination. Haplotype nucleotide sequences were extracted with a small script. Nucleotide sequences were then visualized and mapped to the reference in Geneious v9 (Biomatters) and amino-acid variants were determined. For visualization of the spectra of single-nucleotide variants, NGS reads were aligned to the reference gene. Nucleotide occurrence frequencies were then determined in R on the aligned NGS reads using the deepSNV Bioconductor package[453]. Sequences containing sgRNA/crRNAs from the pooled lentiviral screens were trimmed from adapter sequences. Individual sgRNA/crRNA sequencing reads were then counted with EdgeR[166, 167] and the fold change of enriched sgRNAs was calculated as follows and visualized using GraphPad Prism:

$$\mathrm{Log}_2\left(\frac{1+(\text{Read count per Million Reads})_{after\ drug\ treatment}}{1+(\text{Read count per Million Reads})_{after\ puromycin\ selection}}\right)$$

Cloning of the sgRNA and crRNA libraries

The 2,209 sgRNA sequences used in the ispinesib-KIF11 pilot screen were obtained by selecting all $N_{21}GG$ sequences available in the NCBI consensus coding sequences of the main isoforms of 9 genes (*KIF11*, *XPO1*, *ERCC3*, *PAK4*, *ABL1*, *TUBB*, *ACTB*, *RPS3A*, and *H2BFM2*). Also included were 100 control sgRNAs and all sgRNA sequences were appended 5′ and 3′ with small DNA sequences to facilitate PCR (total length 60 nt). To obtain the AsCpf1 crRNA sequences, the coding sequence of the main isoforms of 10 genes (*KIF11*, *XPO1*, *ERCC3*, *PAK4*, *ABL1*, *TUBB*, *ACTB*, *RPS3A*, *H2BFM2*, and *TP53*) were extracted from NCBI. For each intron–exon boundary, 25 intronic nucleotides were added to the exonic sequences. From these sequences, all $TTTN_{24}$ sequences were extracted and appended 5′ with the AsCpf1 direct repeat backbone (TAATTTCTACTCTTGTAGA) and 30 scrambled control crRNAs were included. Sequences were then further appended 5′ and 3′ with small DNA sequences to facilitate PCR (total length 79 nt).

The sgRNA sequences for the "FDA-target" and "non-FDA-target" libraries were obtained with a custom script run in RStudio. In brief, target genes for the libraries were roughly determined by the drug target list available for approved and investigational antineoplastic agents on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (retrieved July 2016) combined with a small literature study. The target list was then filtered from agents consisting of analogs of nucleotide or metabolic products. The web-based Biomart (Ensemble) was used to obtain the start and end coordinates of all CDS exons retrieved from the NCBI refseq entries available for all isoforms of the predefined target genes. Twenty base pairs were added 5′ and 9 bp were added 3′ to each of the exonic start or end coordinates on the forward and reverse strands, respectively, to include sgRNAs located on exon–intron

boundaries. These expanded, and strand-specific coordinates were used to search through the NCBI reference sequences to obtain all $N_{21}GG$ sequences within these coordinates on both the forward and reverse strands. Duplicate sgRNA sequences were removed on a gene-per-gene basis, sgRNAs containing a TTTTT sequence were removed and sgRNAs were then appended with additional sequences to facilitate PCR and the generation of subpools. See **Supplementary Data 1, 3, 4,** and **8** for the gene target lists and the individual sgRNA sequences.

All appended sgRNA and crRNA sequences were synthesized as pools by Customarray Inc. (Bothell, WA) on a 12K (9/10 gene libraries) or 90K ("non-FDA-target"/"FDA-target") chip. The sgRNA/crRNA pools were amplified in 10 parallel reactions (25 µL, 1 ng input) by PCR with the CloneAmp HiFi premix kit (Clontech) and PCR products were purified with the QIAQuick Nucleotide Removal kit (QiaGen). The purified PCR products were then subjected to restriction digestion (six parallel reactions) with BfuAI (NEB) overnight at 50 °C. After digestion, six ligation reactions containing 33 ng of digested sgRNA/crRNAs and 500 ng of the BfuAI and NsiI predigested pLCKO vector were performed overnight at 16 °C with T4 DNA ligase (NEB). The pooled mixture of ligated pLCKO vectors was then purified with the QIAquick nucleotide removal kit (Qiagen) and electroporated into Endura competent cells (Lucigen) with a Gene Pulser system (Biorad) according to the manufacturer's instructions. Transformed cells were then plated in 15-cm-diameter Petri dishes containing prewarmed LB agar with 100 µg/mL ampicillin and grown overnight at 32 °C. The following day colonies were counted and a fold representation of 400 ("FDA" libraries A and B), 2,700 ("non-FDA" libraries C and D), 30,000 (9 gene Cas9 library), or 90,000 (AsCpf1 library) was estimated. All colonies were pooled per library for plasmid extraction with the PureLink® HiPure Plasmid Maxiprep (Invitrogen).

Lentiviral libraries

The pooled and purified pLCKO-U6-sgRNA/crRNA plasmid libraries were provided to Applied Biological Materials Inc. (Richmond, BC, Canada) to generate lentiviral particles coated with the VSV-G protein and containing the desired genetic information for human expression of the sgRNA/crRNAs. Viral stocks were titrated on wild-type HAP1 or K-562 cells to determine the MOI.

Target identification screens

HAP1 and K-562 cells stably expressing SpCas9 or AsCpf1 were resuspended in supplemented IMDM containing 8 µg/mL polybrene and transduced with lentiviral particles containing the desired sgRNA/crRNAs at an MOI of 0.25 (coverage of 5,000× per sgRNA for SpCas9) or 0.35 (coverage of 15,000× per crRNA for AsCpf1) by spinfection in 12-well plates containing 2×10^6 cells/well. For the bortezomib resistance screen, ~300 million cells were transduced with sublibrary B, while for the KPT-9274 resistance screen ~400 million were transduced with both sublibraries C and D. The next day cells were transferred to T150 cell flasks and were allowed to develop mutations for a period of 3–5 days under puromycin selection. Then 4 million (ispinesib/selinexor) or 10 million (KPT-9274/bortezomib) cells were harvested for DNA extraction and the remaining cells were treated for a period of 2–3 weeks with 8 nM ispinesib, 30 nM bortezomib, 300 nM (HAP1) or 500 nM (K-562) KPT-9274, or 2 µM selinexor. The compound-containing medium was refreshed three times. After treatment, surviving cells were harvested and the genomic DNA was extracted using the QiaGen DNA mini kit and subjected to 5–10 parallel 100 µL PCR reactions (24 cycles, 2,000 ng per reaction) with pLCKO primers carrying Nextera adapter sequences and using the Phusion High-Fidelity PCR mastermix with HF buffer (NEB). Amplified DNA was purified and pooled and a second PCR using the CloneAmp HiFi Polymerase kit (Clontech) was performed over 24 cycles on 50 ng with Nextera indexing primers (Illumina). Further processing for next-generation sequencing analysis was performed as described above.

XPO1-mediated nuclear export phenotypic reporter assay

To study the XPO1-mediated nuclear export, HAP1/SpCas9$^+$ cells were transfected with a plasmid encoding for the NLS-AcGFP-NES reporter construct and a plasmid pool encoding 4 sgRNAs targeting *XPO1* to knockout *XPO1*. Cells were transfected with the Neon Electroporation system as described above and plated in µ-slide 8-well glass bottom plates (IBIDI). Two days after transfection, cells were treated with DMSO or 1 µM selinexor for 3hr. After incubation, the AcGFP reporter construct was visualized in live cells using a Leica SP5 confocal microscope employing a CX PL APO 63x (NA 1.2) objective. AcGFP was excited at 488 nm (argon laser) and emission was detected at 492–560 nm.

Pull-down of XPO1 with KPT-9058

For pull-down of XPO1, parental HAP1 and HDR-edited HAP1 cells were incubated with 2 µM of KPT-9058 for 3hr. Following treatment, cells were washed in ice-cold PBS and pellets were lysed on ice in RIPA buffer supplemented with 1× HALT protease inhibitors (ThermoFisher Scientific). Samples were cleared from debris by centrifugation at 18,000×g for 10 min at 4 °C. Protein concentrations were measured with a colorimetric BSA protein assay (Pierce). A fraction of the cell lysis mixture was taken for quantification of β-tubulin and XPO1 total protein. Remaining extracts were incubated with Dynabeads MyOne Streptavidin T1 (Life Technologies) to capture XPO1-bound KPT-9058 by rotating overnight at 4 °C. The next morning, beads were collected with the DnyaMag-2 (Life Technologies) and washed five times in modified RIPA buffer (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 1% NP-40 (IGEPAL CA-630), 0.1% sodium deoxycholate, 1 mM EDTA). The captured proteins were eluted by boiling the samples for 10 min in 0.5% SDS containing 1× sample buffer (Protein Simple). Samples were separated by size (12–230 kDa) and visualized on a Wes system (Protein Simple) with an anti-rabbit HRP conjugated antibody detecting the primary XPO1 (1/12,500, NB100–79,802) and β-tubulin (1/3,000, NB600–936) antibodies. Protein signals were visualized and quantified with the Compass software, v2.7.1.

Immunofluorescence analysis of mitotic spindle formation

Parental HAP1 and HDR-edited HAP1 cells were electroporated with a plasmid pool encoding for four sgRNAs targeting *KIF11* (for *KIF11* knockout) or the *AAVS1* locus (control) and plated in 8-well chamber glass slides. The next morning, cells were treated with DMSO or 50 nM ispinesib for 4hr. Cells were fixed for 10 min at room temperature using 4% PFA and permeabilized with PBS containing 0.1% Triton X-100 for 10 min. Afterwards, cells were washed with PBS and blocked with 10% normal goat serum for 1hr at 37 °C. Subsequently, cells were stained with an α-tubulin primary antibody (sc5286, Santa Cruz Biotech) dissolved in 10% normal goat serum for 1hr at 37 °C. Following primary staining, cells were washed two times with PBS and stained with a secondary antibody conjugated to Alexa Fluor 488 (1/500, A-11001, Invitrogen) in the dark for 45 min at 37 °C. Nuclei were counterstained using the NucBlue fixed Cell stain (Life Technologies). Stained cells were imaged using a Leica SP5 confocal microscope employing a CX PL APO 63× (NA 1.2) water immersion objective. Alexa Fluor 488 and DAPI were detected using the excitation line of 488 nm (argon laser) or the excitation line of 405 nm (pulsed diode laser), respectively. Blue emission was detected between 410 and 480 nm, and green emission was detected between 492 and 560 nm.

Determination of proteasome activity

Proteasome activity was determined using the Proteasome GloTM Assay (Promega), and the assay was performed according to the manufacturer's instructions. Briefly, site-specific luminogenic substrates were added to determine the activity of the different catalytically active sites (Suc-LLVY-aminoluciferin). Proteasome activity was determined at baseline and after 2hr treatment (at 37 °C, 5% CO_2) with 12.5 nM of bortezomib using a Wallac Victor plate reader.

Determination of NAD+ levels

HAP1 cells were plated in 96-well plates and incubated overnight. The next day, a dilution series of KPT-9274 was added and incubated overnight. NAD+ levels were determined the following day using luminescence and the NAD+/NADH Glo Assay (Promega). NAD+ was obtained from SelleckChem to generate a standard curve.

NAMPT crystallization

NAMPT was purified as previously described[440]. Briefly, full-length human NAMPT was expressed with a C-terminal His tag in *Escherichia coli* at 20°C. A combination of nickel-agarose affinity chromatography, anion exchange and gel-filtration chromatography was used to purify the protein. Crystals were grown for 5 days at 20 °C in hanging drops with a reservoir solution of 0.2 M NaCl, 0.1 M Tris-HCl pH 8.5, and 22% PEG 3350. The drops were composed of 1 µL NAMPT protein (10 mg/mL) and a 5-fold excess of KPT-9274 in a 25 mM Tris-HCl, 150 mM NaCl, 5 mM DTT pH 8.0 buffer and 1 µL of reservoir solution. Crystals were shortly soaked in a cryoprotectant consisting of the reservoir solution with an additional 10% glycerol and flash-frozen in liquid nitrogen. Diffraction data were collected at the SSRF-BL17B beamline at 100 K using 0.95369 Å radiation using a Rayonix MX300 detector and processed with HKL-3000. PDB entry 4WQ6 was used as a starting point for molecular replacement and structure refinement utilizing CCP4[454], Phenix[455], and Coot[456] (see **Supplementary Fig. 11**). The KPT-9247 structure was built into the difference map at the ligand bind site, which shows a good agreement with the electron density (see **Supplementary Fig. 11**). Following introduction of the mutations by modeling utilizing the Molecular Operating Environment (MOE, the Chemical Computing Group, Montreal, Canada) implemented with the Amber99 forcefield[457], the conformation of the residues within a 10Å radius of the ligand was energy optimized and the conformation changes at the ligand-binding site were observed.

Data availability

The coordinates of the crystal are submitted to the online database wwPDB as 5NSD. The used R scripts can be found online in the Supplementary Software file of the associated publication in *Nature Communications*. Sequencing data are available online from the NCBI Sequencing Read Archive under the accession code SRP126384 (Bioproject PRJNA419880).

Acknowledgements

We thank Wim Meert, the KU Leuven Genomics Core and Sarah Gillemot for their technical support with MiSeq next-generation sequencing and Els Vanstreels for help with imaging.

Competing interests

The KU Leuven has filed a patent application based on the findings in this paper. E.B. and W.S. are employees of Karyopharm Therapeutics Inc.

Author contributions

J.E.N. and D.D. designed the study and experiments. J.E.N. conducted experiments with the help of L.B., K.M., and B.M., and analyzed the data. J.E.N. and B.K. designed the CRISPR libraries and B.K. and K.M. cloned the libraries and generated the SpCas9 stable cell lines. T.D. set up the CrispRVariants scripts and wrote other necessary scripts for next-generation sequencing analysis and library design. J.E.N. and T.D. analyzed next-generation sequencing data. M.J. and T.V. cloned the AsCpf1 cell line. N.K. and M.D. performed the PSMB5 protease activity experiments. E.B. and W.S. provided the co-crystal structure and compound, and H.N. and A.V. analyzed and visualized the crystal structure and performed the modeling. J.E.N. designed the figures and J.E.N. and D.D. wrote the manuscript with input from all authors.

Supplementary information

Supplementary Figure 1
XPO1 variants in genome edited and KPT-185 resistant HAP1 cells.

Supplementary Figure 2
KIF11 variants present in genome edited and ispinesib resistant cells.

Supplementary Figure 3
ERCC3 variants present in edited and triptolide resistant cells.

Supplementary Figure 4
Generation of triptolide resistance mutations at the ERCC3$_{D54}$ locus.

Supplementary Figure 5
Assessment of cross-resistance between drug resistant cell pools.

Supplementary Figure 6
Generation of resistant protein variants in HL-60 cells.

Supplementary Figure 7
Generation of drug resistant protein variants in HCT 116 cells.

Supplementary Figure 8
Identification of a drug's target protein from a pool of 9 genes.

Supplementary Figure 9
NAMPT sgRNAs confer resistance to other cancer cells and FK866.

Supplementary Figure 10
Validation of KPT-9274 resistance mutations by Cas9-induced HDR.

Supplementary Figure 11
Details of the cocrystallization of KPT-9274 with NAMPT dimer.

Supplementary Figure 12
AsCpf1 editing of the ERCC3^{D54} locus confers triptolide resistance.

Supplementary Figure 13
Sequencing results of single-cell derived selinexor resistant colonies.

a

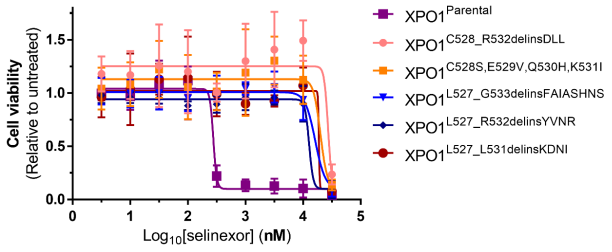
Selection (KPT-185)	Type	Protein Sequence (Cys$_{528}$)	Read Frequency	CrispRVariants Identifier
300 nM	Wild-type	...LGLCEQKRGKD...	1.9%	No variant
300 nM	MNV	. . . SVHI	28.8%	SNV: -10,-7,-6,-3,-1
300 nM	Indel	. . YVNR	4.2%	-14:1D,-2:1I
300 nM	Deletion	. . . - - - -	1.5%	-12:12D
300 nM	Deletion	. . -DPPV	1.2%	-14:1D,-10:1D,-2:1D
300 nM	Frameshift	. . . . . KRQR	27.1%	-2:1I
300 nM	Frameshift	. . . . . EAKI	5.8%	-2:1D
300 nM	Frameshift	. . . . . - -RQR	4.4%	-4:5D
300 nM	Other	-	25.1%	Other
600 nM	Wild-type	...LGLCEQKRGKD...	0.9%	No variant
600 nM	Indel	. . YVNR	36.8%	-14:1D,-2:1I
600 nM	Indel	. . SIRII	5.1%	-13:1I,-5:1D
600 nM	Deletion	. DYVNR - - . . .	2.7%	-17:1D,-4:5D
600 nM	Deletion	. . . - - SI	1.3%	-10:5D,-4:1D
600 nM	Deletion	. . . GL - -	1.3%	-11:1D,-9:1D,-6:4D
600 nM	Frameshift	. . . . . EAKI	37.9%	-2:1D
600 nM	Other	-	14.0%	Other
1500 nM	Wild-type	...LGLCEQKRGKD...	1.1%	no variant
1500 nM	Deletion	. . . - - K	59.1%	-11:4D,-5:1D,-3:1D
1500 nM	Deletion	. . . - -	7.3%	-12:6D
1500 nM	Indel	. . . ILIL	4.6%	-12:1I,-3:1D
1500 nM	MNV	. . . SVHII . . .	2.6%	SNV:-10,-7,-6,-3,-1,3
1500 nM	Other	-	25.3%	Other
2000 nM	Wild-type	...LGLCEQKRG-KD...	1.1%	No variant
2000 nM	Deletion	. . -YVNR	27.5%	-14:1D,-2:2D
2000 nM	Deletion	. . -KDNI	10.1%	-15:2D,-11:1D
2000 nM	Indel	. . FAIASHNS . .	6.2%	-12:2I,-10:1I,-7:1I,5:1D
2000 nM	Deletion	. . -SVHI	1.1%	-14:1D,-8:2D
2000 nM	Frameshift	. . . . . KEA . I	23.1%	-2:2I
2000 nM	Frameshift	. . . . . - -R . QR	11.3%	-4:5D
2000 nM	Frameshift	. . . . . - -A . . I	2.4%	-4:7D
2000 nM	Other	-	17.2%	Other

Selection (KPT-185)	Type	Protein Sequence (Cys$_{528}$)	Read Frequency	CrispRVariants Identifier
300 nM	Wild-type	...LGLCEQKRGKD...	78.0%	No variant
300 nM	Deletion	. . . - - -	1.2%	-12:6D
300 nM	Indel	. . . AAWE	1.1%	-11:1D,-4:1I
300 nM	Frameshift	. . . . . - -QR	3.3%	-2:2D
300 nM	Other	-	16.4%	Other
600 nM	Wild-type	...LGLCEQKRGKD...	9.4%	No variant
600 nM	Deletion	. . . DR - -	12.9%	-11:1D,-9:3D,-2:2D
600 nM	Deletion	. . . SI - -	11.7%	-10:5D,-4:1D
600 nM	Deletion	. . . Q - -	10.7%	-12:6D
600 nM	Deletion	. . SV - -	6.6%	-14:1D,-8:5D
600 nM	Deletion	. . FASI -	5.6%	-12:1D,-9:1D,-4:1D
600 nM	Indel	. . FTGSI	4.7%	-14:1I,-4:1D
600 nM	Deletion	. . KDNI -	1.7%	-15:2D,-11:1D
600 nM	MNV	. . . IDNN	1.5%	SNV:-11,-10,-9,-6,-5,-3,1
600 nM	Deletion	. . . . . - - N . .	1.1%	1:6D
600 nM	Frameshift	. . . . . - -RQR	11.6%	-4:5D
600 nM	Frameshift	. . . . . KRQR	3.0%	-2:1I
600 nM	STOP	. . . * - -	1.3%	-10:9D
600 nM	Other	-	18.2%	Other
1500 nM	Wild-type	...LGLCEQKRGKD...	2.4%	No variant
1500 nM	Indel	. . . SIRI	21.8%	-8:1D,-3:1I
1500 nM	Deletion	. . FVHI -	8.2%	-12:1D,-8:2D
1500 nM	Indel	. . SVNLLG . . .	6.0%	-14:1D,-5:1I
1500 nM	Deletion	. . . K - -	2.2%	-11:4D,-5:1D,-3:1D
1500 nM	Deletion	. . FQR - -	1.3%	-12:1D,-10:5D
1500 nM	Deletion	. . . GL - -	1.2%	-11:1D,-9:1D,-6:4D
1500 nM	Frameshift	. . . . - - - A . I	20.1%	-8:10D
1500 nM	Frameshift	. . . . . - - - I	6.9%	-4:13D
1500 nM	Frameshift	. . . . . - - - R	2.9%	-2:11D
1500 nM	Other	-	27.0%	Other
2000 nM	Wild-type	...LGLCEQ-KRGK...	2.5%	No variant
2000 nM	Insertion	. . FASVHI . . .	23.1%	-12:3I
2000 nM	Indel	. CANRS . I . . .	12.7%	-8:1I,-4:1D
2000 nM	MNV	. CSHNP . N . . .	8.2%	SNV:-17,-15,-13,-11,-10,-8,-6,-4,-3,1
2000 nM	Deletion	. . . Q - -	2.5%	-12:6D
2000 nM	MNV	. . . GKD . LL . .	2.3%	SNV:-11,-9,-8,-5,-3,-2,-1,2,3
2000 nM	Insertion	. . . . . N	1.8%	-2:3I
2000 nM	Deletion	. . SYV - . - . . .	1.4%	-7:6D
2000 nM	Deletion	. . SVHI . - . . .	1.2%	-14:1D,-8:2D
2000 nM	Deletion	. . . SI - . - . . .	1.2%	-10:5D,-4:1D
2000 nM	Frameshift	. . . . . - - RQ	13.8%	-4:5D
2000 nM	Other	-	29.3%	Other

b

Type	Sample (KPT-185)	Apparent Ploidy	*XPO1* Genotype (Cys$_{528}$)	Protein
Wild-type	-	Haploid	ATTAGGATTATGTGAACAGAAAAGAGGC -	- -
Clone 1	300 nM	Haploid	ATTAGGATTAGATCTATT------AGGC -	p.C528_R532delinsDLL -
Clone 2	300 nM	Diploid	ATTAGGATTATCTGTTCACATAAGAGGC ATTAGGATTATGTGAACAGAAA-GAGGC	p.C528S, E529V, Q530H, K531I p.R532fsX16
Clone 3	300 nM	Diploid	ATTAGGATTATCTGTTCACATAAGAGGC ATTAGGATTATGTGAACAGAAA-GAGGC	p.C528S, E529V, Q530H, K531I p.R532fsX16
Clone 4	2,000 nM	Diploid	ATTAGGATTTGCTATTGCCTCACATAATAGC ATTAGGATTATGTGAAC-------AGGC	p.L527_G533delinsFAIASHNS p.K531fsX15
Clone 5	2,000 nM	Diploid	ATTAGGAT-ATGTGAACAG--AAGAGGC ATTAGGATTATGTGAACAGAAGAAAGAGGC	p.L527_R532delinsYVNR p.K531fsX17
Clone 6	2,000 nM	Diploid	ATTAGG--CA-AAGATAATATTAGAGGC ATTAGGATTATGTGAAC-----AGAGGC	p.L527_L531delinsKDNI p.K531fsX4

c



Supplementary Figure 1: *XPO1 variants in genome edited and KPT-185 resistant HAP1 cells.*
a. XPO1 amino-acid variants as detected with CrispRVariants in HAP1 cells transfected with SpCas9 and sgRNA$_{XPO1/L531}$. Cells were treated with four different concentrations of KPT-185 and the experiment was repeated once. Variants detected above a 1% threshold are shown. Cysteine$_{528}$ is highlighted in blue while a red arrowhead and dotted line denote the cut site.
b. Genotypes of single-cell derived, KPT-185 resistant, *XPO1* genome edited HAP1 colonies subjected to Sanger sequencing. The sgRNA targeting sequence is underlined in the reference sequence, while the PAM site is highlighted in red. Bold nucleotides highlight mutations. The codon encoding the cysteine$_{528}$ residue (TGT) is highlighted in the blue column.
c. Cell viability after 72hr of single-cell derived clones obtained from the pool of resistant HAP1 cells in the presence of selinexor. Data points are normalized relative to untreated cells and represent means ± s.d. obtained from two experiments performed in triplicate.

Selection (Ispinesib)	Type	Protein Sequence	Read Frequency	CrispRVariants Identifier
		Asp$_{130}$ ▼ Ala$_{133}$		
4 nM	Wild-type	...DPL-AGIIP...	2.3%	No variant
4 nM	Deletion	--.-.....	51.4%	-9:9D
4 nM	Deletion	...-.....	20.9%	-3:3D
4 nM	Insertion	..FL.....	2.3%	-3:3I
4 nM	Deletion	.--.-....	2.3%	-6:9D
4 nM	Deletion	...-.-....	2.0%	-4:6D
4 nM	Deletion	.-.....	1.3%	-4:3D
4 nM	Deletion	...-.....	1.1%	-2:3D
4 nM	Frameshift	..F.GWYNS	1.4%	-2:1I
4 nM	Frameshift	...--WYNS	1.3%	-3:5D
4 nM	Stop	....L*FQA	1.0%	1:12I
4 nM	Other	-	12.7%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
8 nM	Wild-type	...DP-LAGIIP...	1.0%	No variant
8 nM	Deletion	--.-.....	53.4%	-9:9D
8 nM	Deletion	...-.....	21.1%	-3:3D
8 nM	Insertion	..F.....	2.7%	-3:3I
8 nM	Deletion	...-.....	1.5%	-2:3D
8 nM	Deletion	.-.--....	1.5%	-6:9D
8 nM	Deletion	...--....	1.4%	-4:6D
8 nM	Deletion	.-.....	1.2%	-4:3D
8 nM	Stop	...L*FQAG	1.4%	1:12I
8 nM	Frameshift	...FGWYNS	1.3%	-2:1I
8 nM	Frameshift	...--WYNS	1.0%	-3:5D
8 nM	Other	-	12.5%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
20 nM	Wild-type	...DPLAGIIP...	0.0%	No variant
20 nM	Deletion	...-.....	39.5%	-3:3D
20 nM	Deletion	--.-.....	28.9%	-9:9D
20 nM	Deletion	...---..	1.8%	-1:9D
20 nM	Deletion	E-.....	1.7%	-6:6D
20 nM	Deletion	...--....	1.7%	-4:6D
20 nM	Deletion	...-.....	1.5%	-2:3D
20 nM	Deletion	.---....	1.3%	-6:9D
20 nM	Frameshift	..FGWYNS	3.8%	-2:1I
20 nM	Frameshift	...--WYNS	1.7%	-3:5D
20 nM	Frameshift	....GWYNS	1.5%	1:1I
20 nM	Frameshift	...-WLV*F	1.5%	-2:1D
20 nM	Frameshift	---WLV*F	1.2%	-14:13D
20 nM	Frameshift	E-----F	1.2%	-7:16D
20 nM	Other	-	12.7%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
40 nM	Wild-type	...DPL-AGIIP...	0.9%	No variant
40 nM	Deletion	--.-.....	43.6%	-9:9D
40 nM	Deletion	...-.....	25.7%	-3:3D
40 nM	Deletion	.-.....	4.7%	-4:3D
40 nM	Deletion	...-.-....	2.3%	-4:6D
40 nM	Deletion	-A.....	1.5%	-7:3D
40 nM	Insertion	..FL.....	1.5%	-3:3I
40 nM	Deletion	...-.....	1.2%	-5:1D,-1:2D
40 nM	Deletion	.--.-....	1.1%	-6:9D
40 nM	Deletion	E-.....	1.0%	-6:6D
40 nM	Frameshift	...FGWYNS	2.0%	-2:1I
40 nM	Frameshift	.--.---NS	1.9%	-7:14D
40 nM	Other	-	12.6%	Other

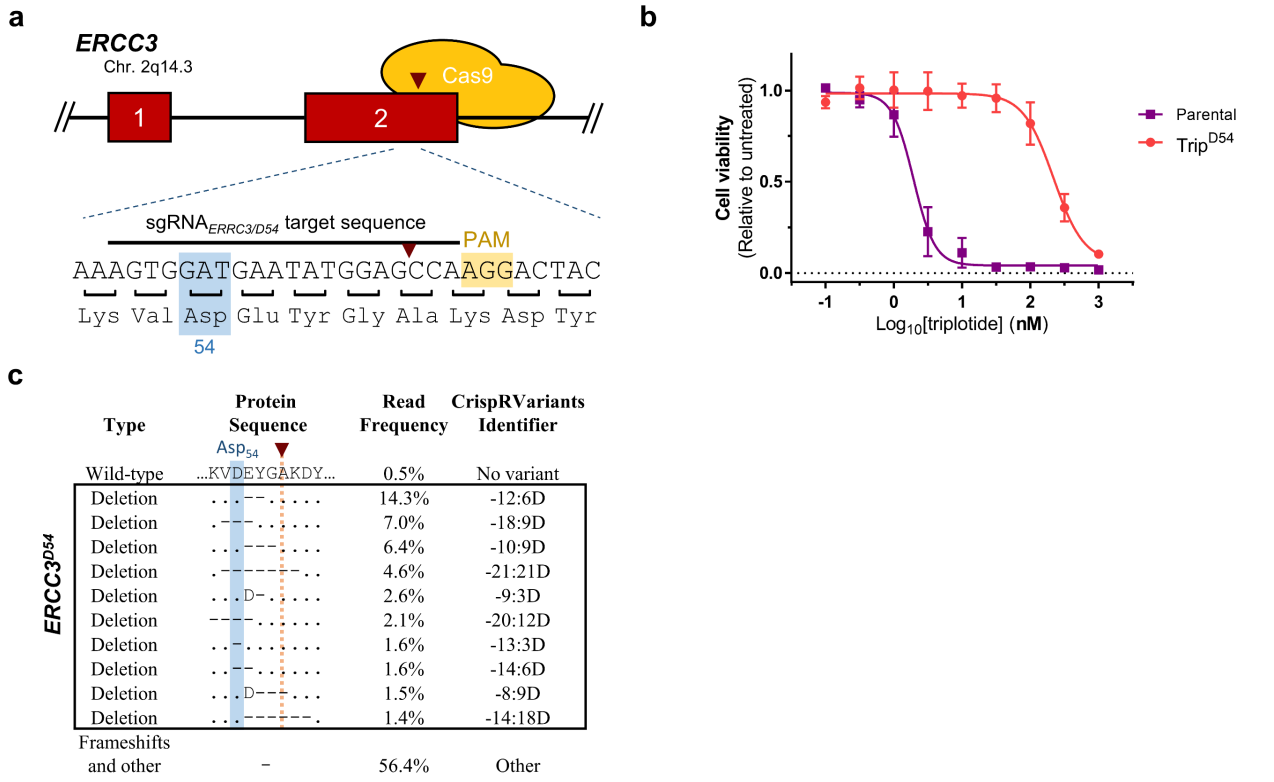
Selection (Ispinesib)	Type	Protein Sequence	Read Frequency	CrispRVariants Identifier
		Asp$_{130}$ ▼ Ala$_{133}$		
4 nM	Wild-type	...DP-LAGIIP...	1.1%	No variant
4 nM	Deletion	--.-.....	53.4%	-9:9D
4 nM	Deletion	...-.....	18.9%	-3:3D
4 nM	Deletion	.X.-.....	2.7%	-4:3D
4 nM	Insertion	..F.....	2.2%	-3:3I
4 nM	Deletion	...--....	1.5%	-4:6D
4 nM	Deletion	.-.--....	1.1%	-6:9D
4 nM	Frameshift	...FGWYNS	2.8%	-2:1I
4 nM	Frameshift	....GWYNS	1.2%	1:1I
4 nM	Frameshift	...FWLV*F	1.1%	-2:2I
4 nM	Frameshift	...--WYNS	1.0%	-3:5D
4 nM	Other	-	13.0%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
8 nM	Wild-type	...DP-LAGIIP...	0.3%	No variant
8 nM	Deletion	--.-.....	54.1%	-9:9D
8 nM	Deletion	...-.....	17.9%	-3:3D
8 nM	Deletion	.X.-.....	3.7%	-4:3D
8 nM	Insertion	..F.....	2.3%	-3:3I
8 nM	Deletion	.-.--....	1.6%	-6:9D
8 nM	Deletion	E-.-.....	1.3%	-6:6D
8 nM	Deletion	...--....	1.1%	-4:6D
8 nM	Frameshift	...FGWYNS	2.4%	-2:1I
8 nM	Frameshift	....GWYNS	1.4%	1:1I
8 nM	Frameshift	--.-GWYNS	1.0%	-8:8D
8 nM	Frameshift	...--WYNS	1.0%	-3:5D
8 nM	Other	-	11.9%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
20 nM	Wild-type	...DP-LAGIIP...	0.2%	No variant
20 nM	Deletion	--.-.....	50.7%	-9:9D
20 nM	Deletion	...-.....	20.5%	-3:3D
20 nM	Deletion	.X.-.....	3.3%	-4:3D
20 nM	Deletion	...--....	2.8%	-4:6D
20 nM	Insertion	..F.....	2.5%	-3:3I
20 nM	Deletion	.-.--....	1.6%	-6:9D
20 nM	Deletion	E-.-.....	1.3%	-6:6D
20 nM	Frameshift	...FGWYNS	2.4%	-2:1I
20 nM	Frameshift	....GWYNS	1.8%	1:1I
20 nM	Frameshift	....-LV*F	1.6%	1:1D
20 nM	Other	-	11.3%	Other
		Asp$_{130}$ ▼ Ala$_{133}$		
40 nM	Wild-type	...DP-LAGIIP...	0.4%	No variant
40 nM	Deletion	--.-.....	38.4%	-9:9D
40 nM	Deletion	...-.....	26.9%	-3:3D
40 nM	Deletion	.-.--....	4.3%	-6:9D
40 nM	Deletion	.X.-.....	3.6%	-4:3D
40 nM	Insertion	..F.....	2.9%	-3:3I
40 nM	Deletion	...-.....	2.7%	-2:3D
40 nM	Deletion	...--....	2.0%	-4:6D
40 nM	Deletion	....---..	1.0%	-1:9D
40 nM	Deletion	...-.....	1.0%	-5:1D,-1:2D
40 nM	Frameshift	...FGWYNS	2.2%	-2:1I
40 nM	Frameshift	....GWYNS	2.0%	1:1I
40 nM	Frameshift	...FWLV*F	1.0%	-2:2I
40 nM	Other	-	11.4%	Other

Supplementary Figure 2: *KIF11 variants present in genome edited and ispinesib resistant cells.* KIF11 amino-acid variants detected in HAP1 cells transfected with SpCas9 and sgRNA$_{KIF11/L132}$. Transfected cells were treated with four different concentrations of ispinesib and the experiment was repeated twice, but the results from only 2 experiments are shown. Protein variants were obtained by CrispRVariants analysis of next generation sequencing reads. Only variants detected above the 1% threshold are shown. Residues known to provide drug resistance upon mutation are highlighted in blue and the Cas9 cut site is shown with a red arrowhead and a vertical dotted line.

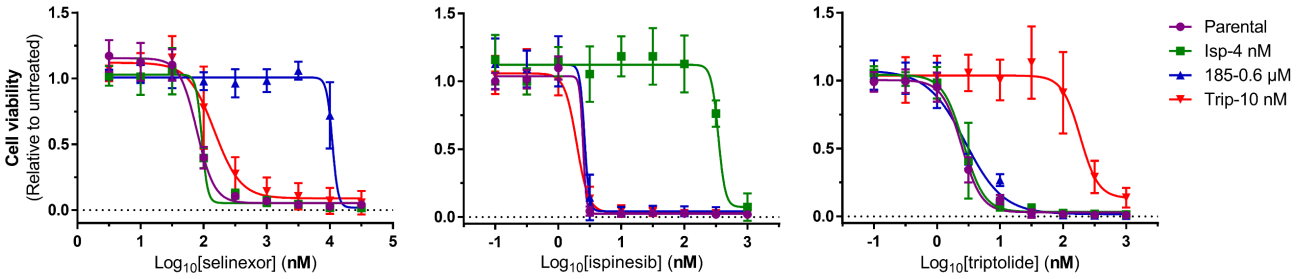
Selection (Triptolide)	Type	Protein Sequence	Read Frequency	CrispRVariants Identifier
		Ser$_{162}$/Tyr$_{163}$ ▼		
2 nM	Wild-type	...VSYGKV-KLVL...	81.8%	No variant
2 nM	Insertion	.....L....	3.1%	-1:3I
2 nM	MNV	.....D.Q...	1.1%	SNV:-3,-1
2 nM	Frameshift	.....-.-AG.	2.3%	-6:5D
2 nM	Frameshift	.......AG.	2.2%	-1:1I
2 nM	Other	-	9.5%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
4 nM	Wild-type	...VSYGKV--KLVL...	21.1%	No variant
4 nM	Insertion	.....L....	24.6%	-1:3I
4 nM	Deletion	.........--.	3.9%	-4:9D
4 nM	MNV	.....D..Q...	3.7%	SNV:-3,-1
4 nM	Insertion	.....AME....	2.9%	-3:1I,-2:5I
4 nM	Insertion	.....XQ...	1.7%	1:3I
4 nM	Deletion	.....-..-...	1.5%	-7:6D
4 nM	Frameshift	.....---.AG.	8.1%	-6:5D
4 nM	Frameshift	.........AG.	7.2%	-1:1I
4 nM	Frameshift	..----...-WS	2.7%	-16:16D
4 nM	Frameshift	.........XAG.	2.7%	1:1I
4 nM	Frameshift	.........-WS	1.2%	-2:4D
4 nM	Other	-	18.7%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
10 nM	Wild-type	...VSYGKV--KLVL...	0.9%	No variant
10 nM	Insertion	.....L....	41.1%	-1:3I
10 nM	Insertion	.....XQ...	3.5%	1:3I
10 nM	Deletion	.....-..-...	2.1%	-7:6D
10 nM	Insertion	.....X....	2.1%	-2:3I
10 nM	Insertion	.....AME....	1.7%	-3:1I,-2:5I
10 nM	Deletion	.........--.	1.5%	-4:9D
10 nM	Insertion	.....-VV....	1.4%	-5:3I
10 nM	SNV	....N.....	1.1%	SNV:-5
10 nM	Insertion	.....TXQ...	1.0%	1:6I
10 nM	Frameshift	.........AG.	11.6%	-1:1I
10 nM	Frameshift	.....-...-AG.	11.3%	-6:5D
10 nM	Frameshift	..-------.-WS	3.6%	-16:16D
10 nM	Frameshift	....-------.	2.0%	-11:14D
10 nM	Other	-	15.1%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
20 nM	Wild-type	...VSYGKV--KLVL...	1.6%	No variant
20 nM	Insertion	.....L....	39.0%	-1:3I
20 nM	Insertion	.....XQ...	9.5%	1:3I
20 nM	Deletion	.........--.	2.1%	-4:9D
20 nM	Insertion	.....X....	1.9%	-2:3I
20 nM	Insertion	.....TL....	1.6%	-1:6I
20 nM	Insertion	.....AME....	1.5%	-3:1I,-2:5I
20 nM	Frameshift	.........AG.	9.3%	-1:1I
20 nM	Frameshift	.........-AG.	9.0%	-6:5D
20 nM	Frameshift	..----...-WS	4.2%	-16:16D
20 nM	Frameshift	.........-SWS	2.2%	-1:1D
20 nM	Frameshift	.........XAG.	1.8%	1:1I
20 nM	Frameshift	.........-WS	1.3%	-2:4D
20 nM	Frameshift	.........--T	1.0%	1:7D
20 nM	Frameshift	.........----	1.0%	-1:10D
20 nM	Other	-	13.0%	Other

Selection (Triptolide)	Type	Protein Sequence	Read Frequency	CrispRVariants Identifier
		Ser$_{162}$/Tyr$_{163}$ ▼		
2 nM	Wild-type	VSYGKV-KLVL	82.4%	No variant
2 nM	Insertion	.....L....	4.5%	-1:3I
2 nM	MNV	.....D-Q...	1.6%	SNV:-3,-1
2 nM	Deletion	.....-.---.	1.0%	-4:9D
2 nM	Frameshift	.....---AG.	2.7%	-6:5D
2 nM	Frameshift	.....-.AG.	2.3%	-1:1I
2 nM	Other	-	5.5%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
4 nM	Wild-type	VSYGKV--KLVL	14.0%	No variant
4 nM	Insertion	.....-L....	31.5%	-1:3I
4 nM	Deletion	.....-----.	6.2%	-4:9D
4 nM	MNV	.....D--Q...	3.6%	SNV:-3,-1
4 nM	Insertion	.....-X....	2.9%	-2:3I
4 nM	Insertion	.....-TQ...	2.1%	1:3I
4 nM	Deletion	.....---M..	1.4%	1:3D
4 nM	Insertion	.....AME....	1.1%	-3:1I,-2:5I
4 nM	Insertion	.....KTT...	1.1%	2:6I
4 nM	Frameshift	.....----AG.	11.2%	-6:5D
4 nM	Frameshift	.....--.AG.	6.0%	-1:1I
4 nM	Frameshift	..-------WS	2.0%	-16:16D
4 nM	Frameshift	....-------.	1.7%	-11:14D
4 nM	Frameshift	.....--RAG.	1.3%	1:1I
4 nM	Frameshift	.....----WS	1.1%	-2:4D
4 nM	Frameshift	.....---SWS	1.0%	-1:1D
4 nM	Other	-	11.8%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
10 nM	Wild-type	VSYGKV-----KLVL	0.6%	No Variant
10 nM	Insertion	.....----L....	39.1%	-1:3I
10 nM	Insertion	.....----TQ...	5.0%	1:3I
10 nM	Deletion	.....--------.	2.7%	-4:9D
10 nM	Insertion	.....----X....	2.6%	-2:3I
10 nM	Insertion	.....AME---....	1.9%	-3:1I,-2:5I
10 nM	Insertion	.....----I....	1.6%	-3:3I
10 nM	Insertion	.....L----FQ...	1.5%	-4:3I
10 nM	Insertion	.....KLWKSQ...	1.1%	1:15I
10 nM	Insertion	.....--XXX....	1.0%	-1:9I
10 nM	Frameshift	.....--------AG.	8.5%	-6:5D
10 nM	Frameshift	.....------.AG.	7.1%	-1:1I
10 nM	Frameshift	..----------WS	4.0%	-16:16D
10 nM	Frameshift	.....------SWS	3.9%	-1:1D
10 nM	Frameshift	....----------.	3.1%	-11:14D
10 nM	Frameshift	.....-----RAG.	1.7%	1:1I
10 nM	Frameshift	.....-----TSWS	1.6%	-2:2I
10 nM	Stop	..---------*...	1.3%	-12:12D
10 nM	Frameshift	....----------	1.2%	-4:3D,2:10D
10 nM	Other	-	10.5%	Other
		Ser$_{162}$/Tyr$_{163}$ ▼		
20 nM	Wild-type	VSYGKV---KLVL	0.0%	No variant
20 nM	Insertion	.....--L....	37.6%	-1:3I
20 nM	Insertion	.....--I....	4.5%	-3:3I
20 nM	Deletion	.....----....	3.8%	-7:6D
20 nM	Insertion	.....--MR...	2.4%	1:2I,2:1I
20 nM	Insertion	.....A--GQ...	2.1%	-3:1I,-2:1I,1:1I
20 nM	Insertion	.....--TQ...	2.0%	1:3I
20 nM	Insertion	.....--X....	1.8%	-2:3I
20 nM	Insertion	.....-KF....	1.8%	-3:6I
20 nM	Insertion	....IWKSL....	1.7%	-13:7I,-1:2I
20 nM	Deletion	.....------M	1.6%	1:9D
20 nM	Insertion	.....-XX....	1.3%	-1:6I
20 nM	Frameshift & other	-	39.4%	Other

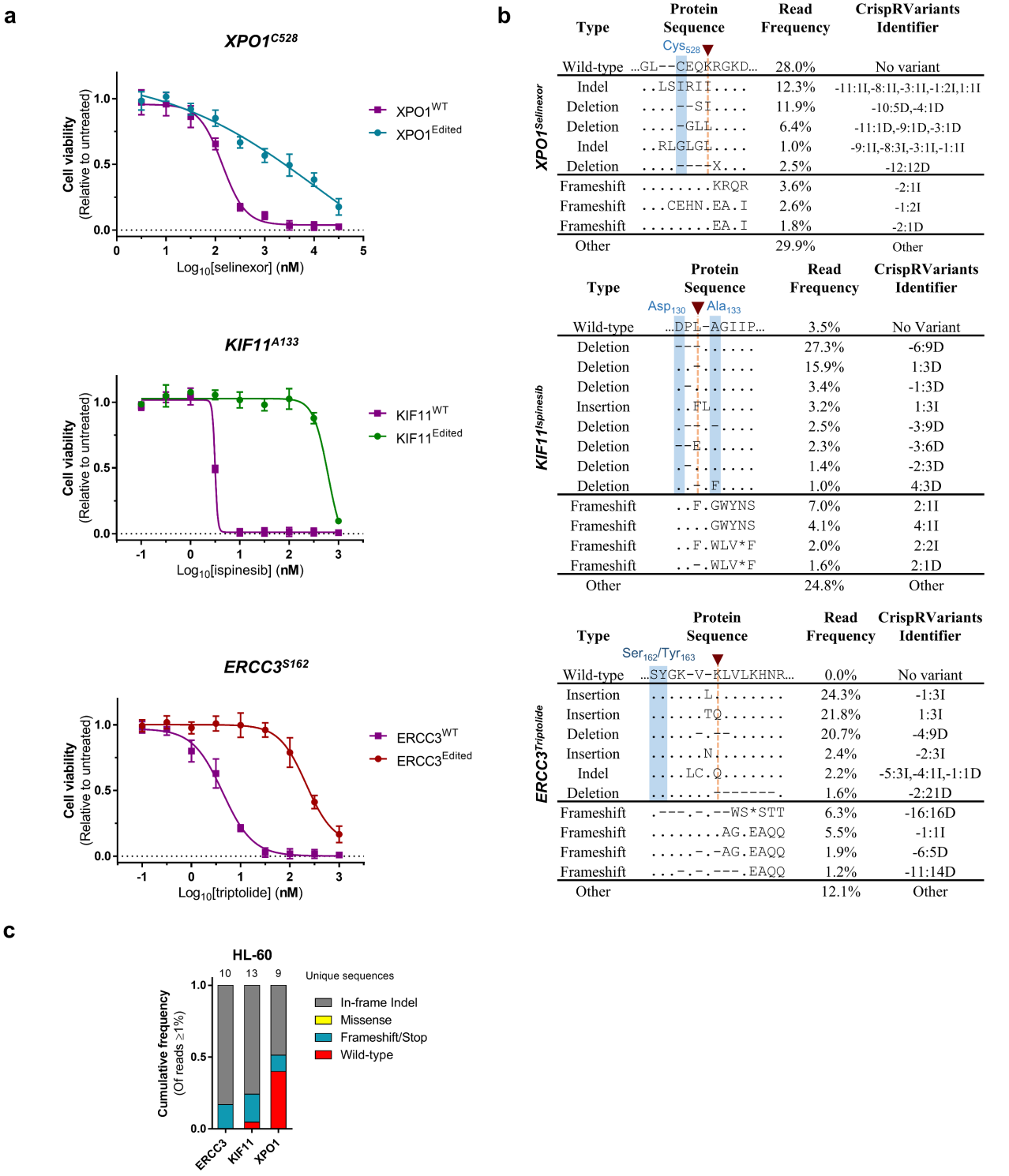
Supplementary Figure 3: *ERCC3 variants present in edited and triptolide resistant cells.*
ERCC3 protein variants detected in HAP1 cells transfected with SpCas9 and sgRNA$_{ERCC3/K167}$. Transfected cells were treated with four different concentrations of triptolide and the experiment was repeated once. Protein variants were obtained by CrispRVariants analysis of next generation sequencing reads. Only variants detected above the 1% threshold are shown. Residues known to provide drug resistance upon mutation are highlighted in blue and the Cas9 cut site is shown with a red arrowhead and a vertical dotted line.



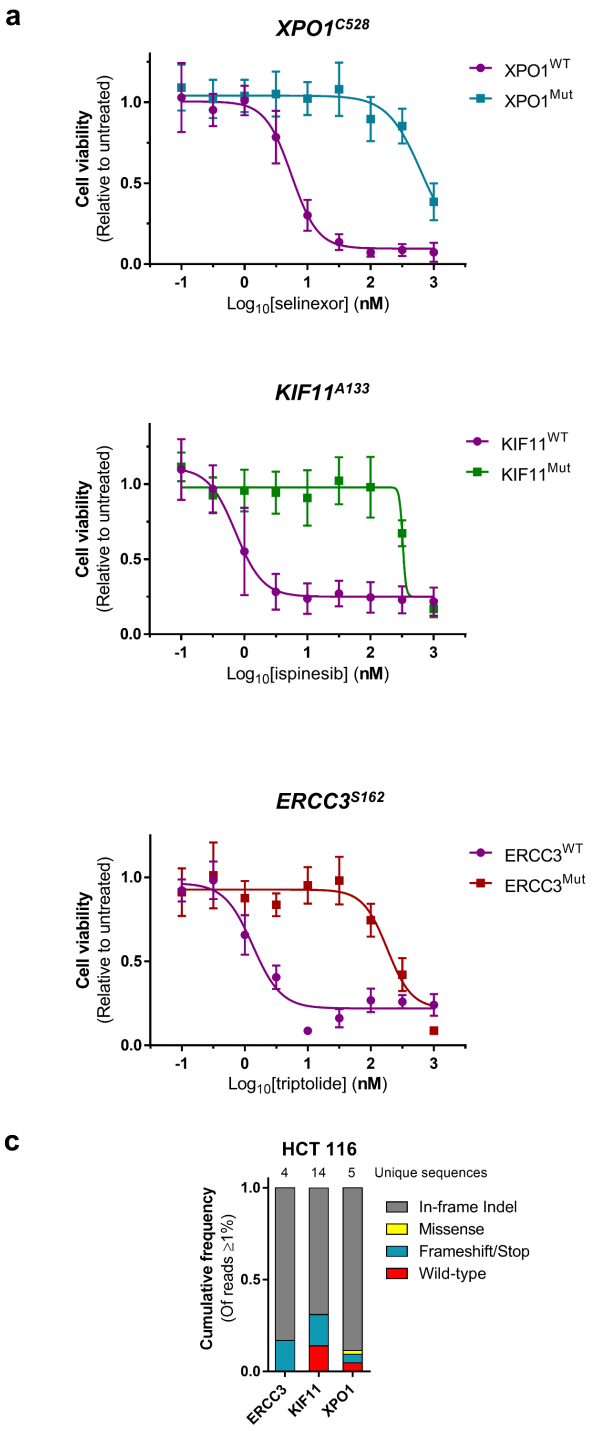
Supplementary Figure 4: *generation of triptolide resistance mutations at the ERCC3$_{D54}$ locus.*
a. Schematic overview of the sgRNA targeting close to the *ERCC3* aspartic acid$_{54}$ *codon*, used for generation of triptolide resistance through CRISPR-SpCas9-mediated genome editing.
b. Cell viability assay showing the effect of triptolide on control and triptolide resistant, *ERCC3$_{D54}$* mutagenized HAP1 cells. Data points are normalized to untreated cells and represent means ± s.d. obtained from three experiments performed in triplicates. Cells were treated for 72h.
c. Amino-acid sequences of SpCas9/sgRNA$_{ERCC3/D54}$ transfected HAP1 cells treated with 10 nM triptolide as determined by CrispRVariants analysis of targeted amplicon sequencing reads. The SpCas9 cut site is shown with a red arrowhead and a vertical dotted line, the aspartic acid$_{54}$ is highlighted with a blue column.



Supplementary Figure 5: *assessment of cross-resistance between drug resistant cell pools.*
Cell viability assays showing the effect of selinexor, ispinesib or triptolide on wild-type (parental) and CRISPR/SpCas9-mutagenized HAP1 cells obtained in **Fig. 1**. For each drug-resistant drug-target pair, cross-resistance against the other two compounds was assessed. For each drug, only a single genome-edited polyclonal mixture of resistant cells was selected for the tests (KPT-185: 600 nM, ispinesib 4 nM, triptolide 10 nM). Data points are normalized relative to untreated cells and represent averages ± s.d. obtained from three experiments performed in triplicate. Cells were treated for 72hr.



Supplementary Figure 6: *generation of drug resistant protein variants in HL-60 cells.*
Mutagenized cells were treated for 2 weeks with 1 µM selinexor (*XPO1*), 10 nM ispinesib (*KIF11*) or 20 nM triptolide (*ERCC3*). SgRNAs as shown in **Fig. 1a** were used for mutagenesis.
a. Cell viability after 72hr of parental and drug resistant SpCas9-mutagenized HL-60 cells in the presence of increasing concentrations of the respective drugs. Data points are normalized to untreated cells and represent means ± s.d. from two experiments performed in triplicate.
b. Amino-acid sequences of SpCas9 mutagenized and drug resistant HL-60 cells as determined by CrispRVariants analysis of targeted amplicon sequencing. Residues known to provide drug resistance upon mutation are highlighted with blue columns. The Cas9 cut site is highlighted by the orange, vertical dotted line. Only variants with a read frequency ≥1% are shown.
c. Overview of the abundance of the different mutation types in the mutagenized HL-60 cells. Only 1 experiment was performed. (*ERCC3*: locus S162, *KIF11*: locus A133, *XPO1*: locus C528).



b

XPO1Selinexor

Type	Protein Sequence (Cys528)	Read Frequency	CrispRVariants Identifier
Wild-type	...LGLCEQ-KRG...	3.1%	No variant
Deletion	...---.....	57.0%	-12:6D
MNV	...SLP.LL.	1.4%	SNV:-10,-8,-7,-6,-4,-3,-2,-1,2,3
Frameshift	.....K..Q	3.0%	-2:1I
Other		35.5%	Other

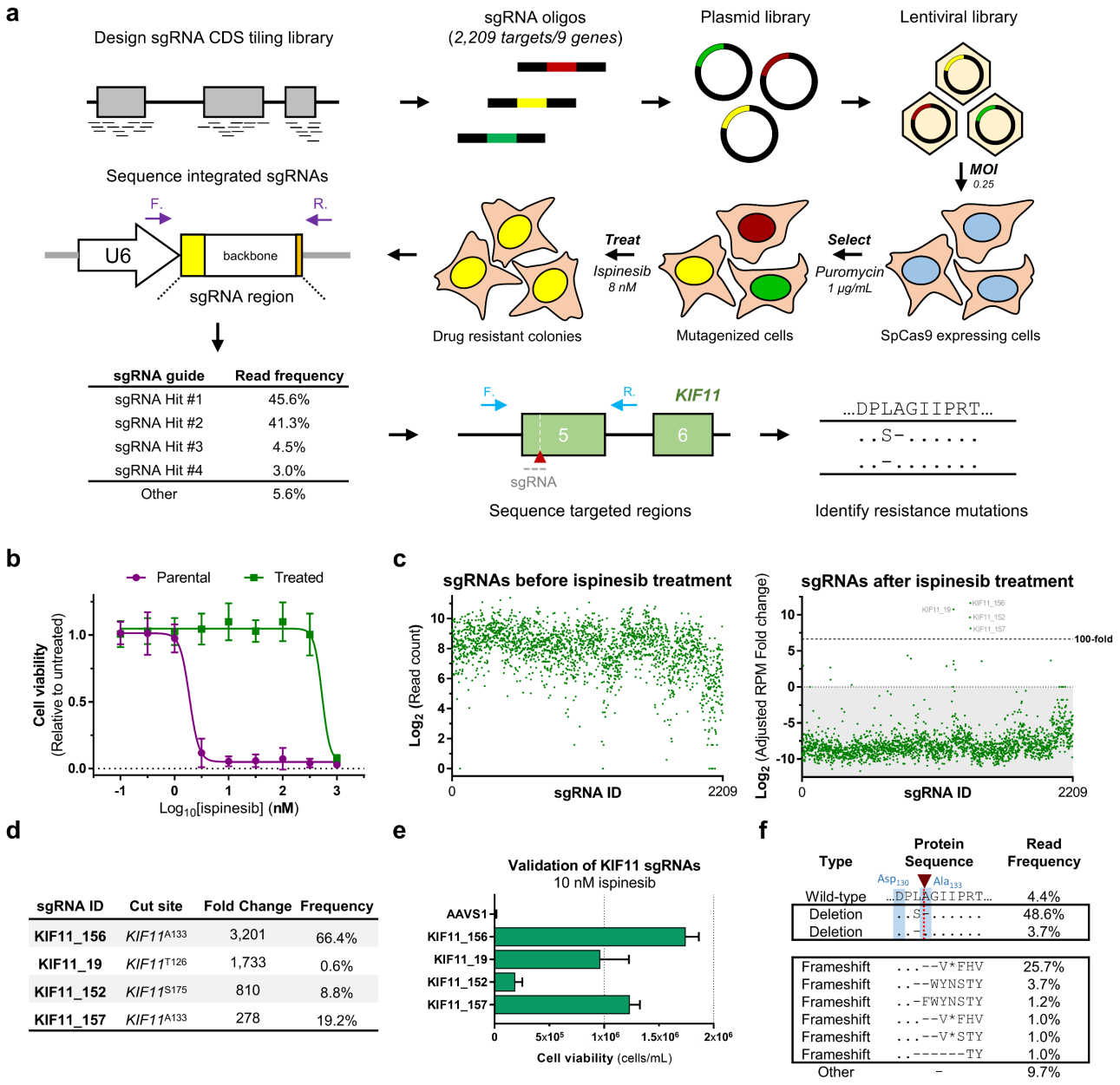
KIF11Ispinesib

Type	Protein Sequence (Asp130 Ala133)	Read Frequency	CrispRVariants Identifier
Wild-type	...DPLAGIIP...	11.3%	No variant
Deletion	---.....	25.2%	-9:9D
Deletion	.-P.....	19.1%	-3:3D
Deletion	.-.....	6.4%	-4:3D
Deletion	.-.....	1.5%	-6:3D
Deletion	.-.....	1.4%	-5:3D
Deletion	--E.....	1.1%	-6:6D
Deletion	A-.....	1.0%	-7:3D
Frameshift	..FGWYNS	6.3%	-2:1I
Frameshift	..FWLV*F	2.0%	-2:2I
Frameshift	..-WLV*F	1.6%	-2:1D
Frameshift	..XGWYNS	1.5%	1:1I
Frameshift	..-WLV*F	1.2%	-5:1D
Frameshift	.-----NS	1.2%	-7:14D
Other		19.2%	Other

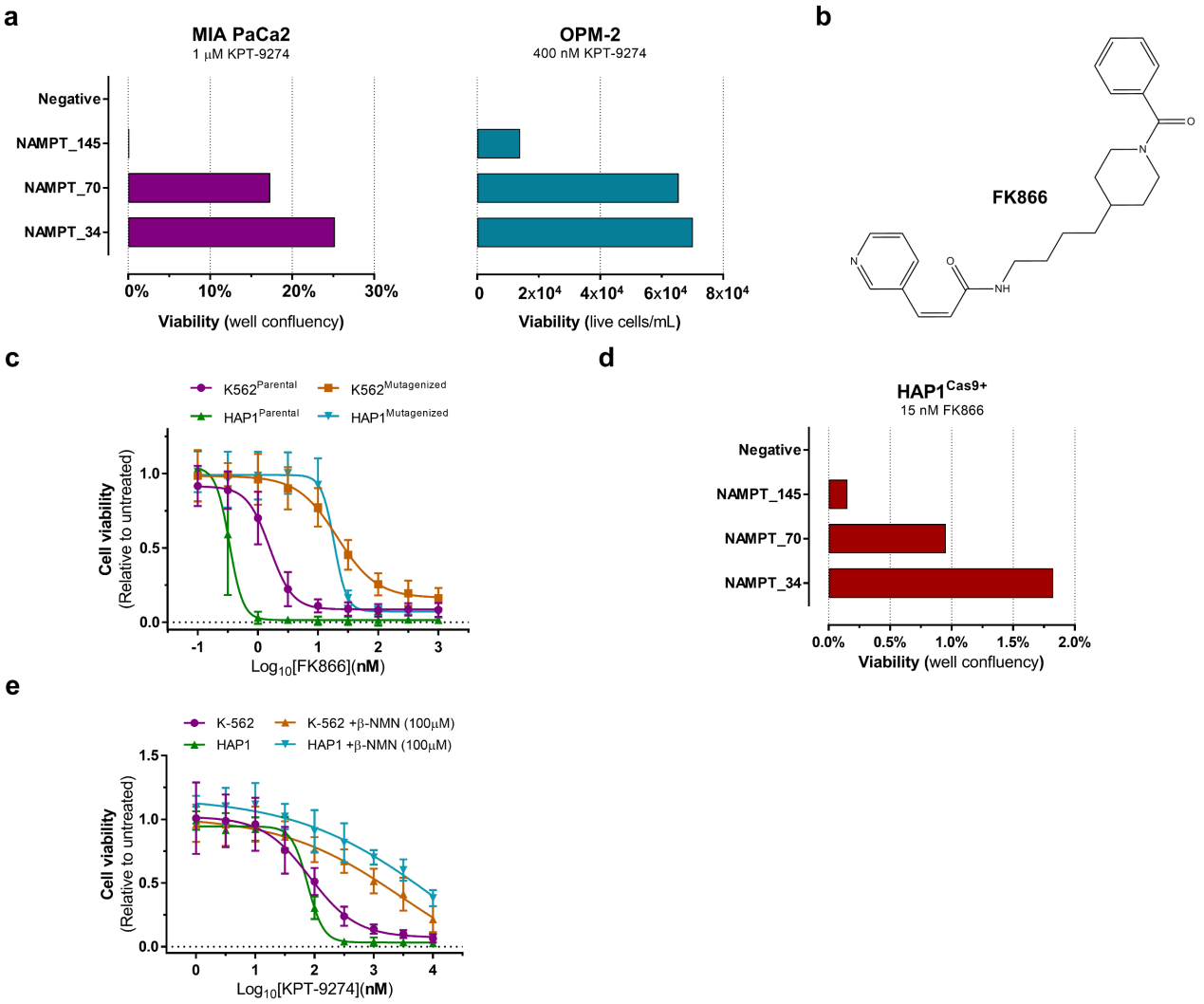
ERCC3Triptolide

Type	Protein Sequence (Ser162/Tyr163)	Read Frequency	CrispRVariants Identifier
Wild-type	...SYGKV--KLVLKH...	0.0%	No Variant
Insertion	.....L.....	41.5%	-1:3I
Indel	...AME.....	39.6%	-3:1I,-2:5I
Frameshift	...---.---.EA	8.8%	-11:14D
Frameshift	...---.----.V	7.7%	-2:17D
Other		2.4%	Other

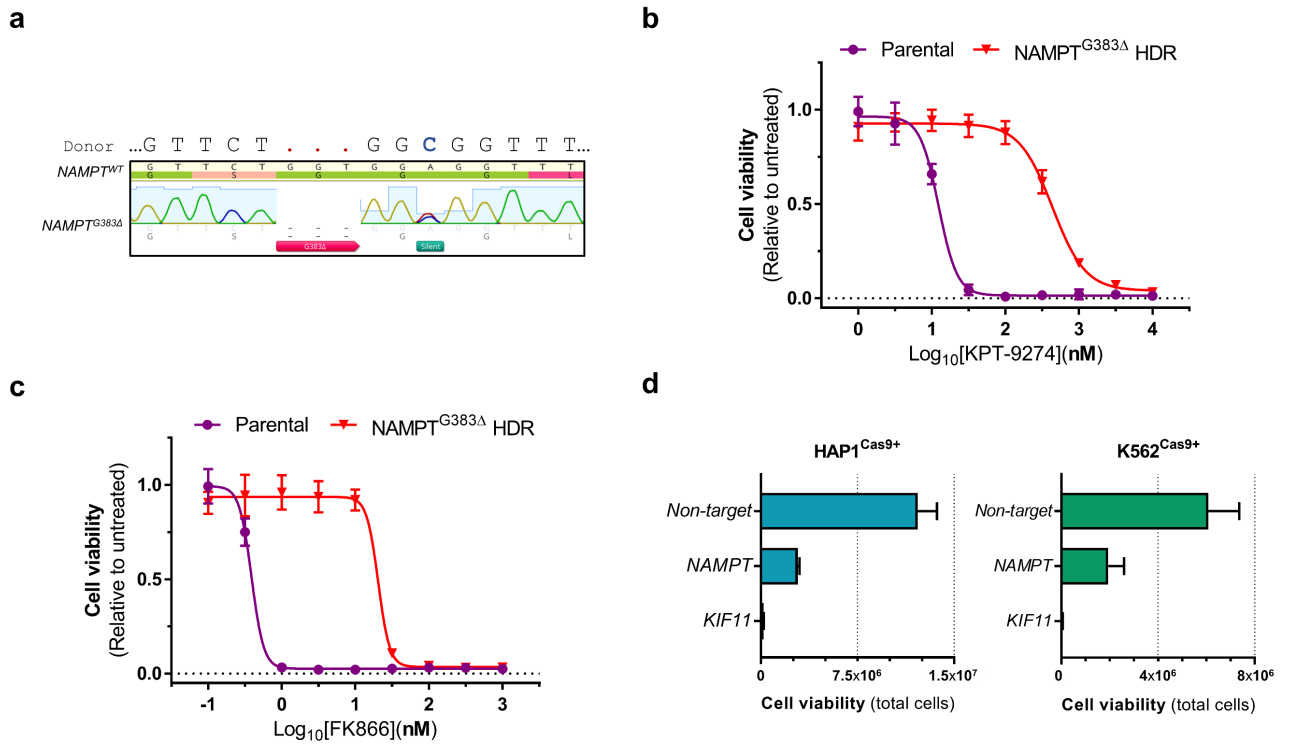
Supplementary Figure 7: *generation of drug resistant protein variants in HCT 116 cells.*
Transfected cells were treated for 2 weeks with 1 µM selinexor (*XPO1*), 10 nM ispinesib (*KIF11*) or 20 nM triptolide (*ERCC3*). SgRNAs as shown in **Fig. 1a** were used for mutagenesis.
a. Cell viability after 72hr of parental and drug resistant SpCas9 mutagenized HCT 116 cells in the presence of increasing concentrations of the respective drugs. Data points are normalized to untreated cells and represent means ± s.d. from three experiments performed in triplicate.
b. Amino-acid sequences of the corresponding SpCas9 mutagenized and drug resistant HCT 116 cells as determined by CrispRVariants analysis. Residues known to provide drug resistance upon mutation are highlighted with blue columns. The Cas9 cut site is highlighted by the vertical dotted line. Only variants detected with a read frequency ≥1% are shown.
c. Overview of the abundance of the mutation types detected in drug resistant HCT 116 cells. Only 1 experiment was performed. (*ERCC3*: locus S162, *KIF11*: locus A133, *XPO1*: locus C528).



Supplementary Figure 8: *identification of a drug's target protein from a pool of 9 genes.*
a. Overview of the CRISPR/Cas9-based target identification screen used for ispinesib.
b. Cell viability after 72hr of parental, Cas9+ and the polyclonal pool of drug resistant mutagenized HAP1 cells treated with different concentrations of ispinesib. Data points represent means ± s.d. obtained from two experiments performed in triplicate.
c. Representation of the sgRNAs in cells before (after puromycin selection) and after treatment with 8 nM ispinesib as determined by EdgeR analysis of next generation sequencing data. Each dot represents a different sgRNA.
d. Overview of enriched sgRNAs with a fold change >100 in ispinesib-resistant cells.
e. Enriched sgRNAs were cloned individually and transfected into parental HAP1 cells stably expressing SpCas9 to validate the results. Three days after transfection, cells were treated with 10 nM ispinesib for 5 days and surviving cells were counted using trypan blue exclusion. An sgRNA targeting the *AAVS1* safe harbor locus on chromosome 19q13.42 was used as negative control. Values represent means ± s.d. obtained from two independent experiments.
f. Amino-acid variants detected in the *KIF11* A133 locus of the resistant pool of cells. The D130 and the A133 residue are highlighted and the sgRNA cut site (KIF11_156/157) is highlighted with a red arrowhead. Alterations were uncovered by CrispRVariants analysis of next generation sequencing data. Only variants with a frequency above 1% are shown.

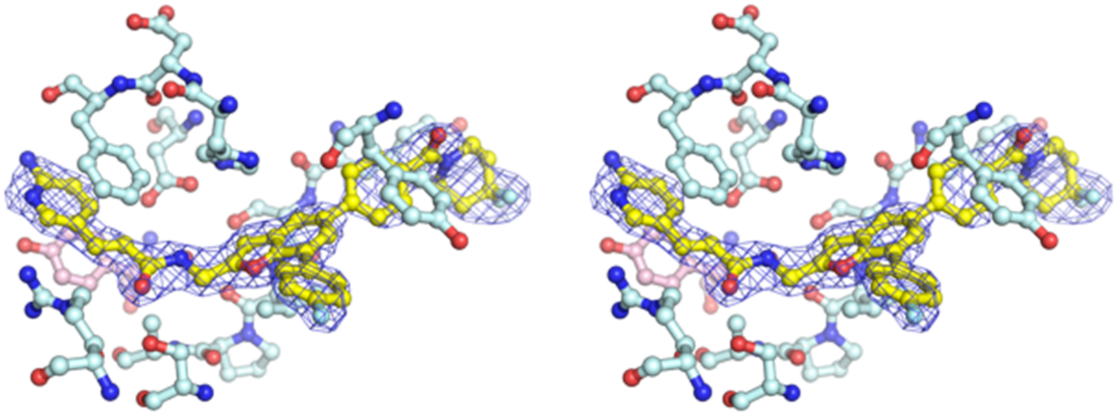


Supplementary Figure 9: *NAMPT sgRNAs confer resistance to other cancer cells and FK866.*
a. Selection of KPT-9274 resistance in MIA PaCa2 and OPM-2 cells using *NAMPT* targeting sgRNAs identified in the KPT-9274 screen with HAP1 and K-562 cells (**Fig. 3d-f**). Cells were cotransfected with the indicated *NAMPT* targeting sgRNAs and SpCas9. Two to three days after transfection, cells were treated for 8 days (MIA PaCa2) or 22 days (OPM-2) with the indicated concentration of KPT-9274. For MIA PaCa2 cells, confluency was measured using an IncuCyte Zoom. For OPM-2 cells, surviving cells were counted by trypan blue exclusion.
b. Chemical structure of the NAMPT inhibitor FK866.
c. The cell populations that survived the mutagenesis screen with KPT-9274 are cross-resistant to FK866. Cell viability assays show the effect of FK866 on parental HAP1 cells stably expressing SpCas9 and the polyclonal mutagenized KPT-9274 resistant cells obtained from the CRISPR/Cas scanning screen. Data points were obtained 72hr after addition of the compound, were normalized relative to untreated cells and represent averages ± s.d. obtained from two (HAP1) or five (K-562) experiments performed in triplicate.
d. *NAMPT* targeting sgRNAs enriched in the KPT-9274 mutagenesis screen confer FK866 resistance to HAP1 cells. Parental HAP1 cells stably expressing SpCas9 were transfected with the indicated sgRNAs and treated with 15 nM FK866 for a period of 8 days. Cell confluency was then measured using an IncuCyte Zoom.
e. Cell viability assay showing the protective effect of β-nicotinamide mononucleotide (β-NMN) on parental HAP1 cells stably expressing SpCas9 in the presence of increasing concentrations of KPT-9274. Data points were obtained after 72h, normalized to untreated cells and represent averages ± s.d. obtained from two (HAP1) or four (K-562) experiments performed in triplicate.



Supplementary Figure 10: *validation of KPT-9274 resistance mutations by Cas9-induced HDR.*
a. Sequencing chromatogram of *NAMPT* in HDR-edited HAP1 cells stably expressing SpCas9. Cells were transfected with a *NAMPT* targeting sgRNA and an HDR donor template containing the G383del mutation together with a silent mutation not identified in our screen to control for HDR. Two days after transfection, cells were treated with 300 nM KPT-9274 for 1-2 weeks.
b. Cell viability assay showing the effect of KPT-9274 on wild-type, Cas9+ (parental) and polyclonal G383del HDR-edited HAP1 cells after 72hr. Data points are normalized to untreated cells and represent averages ± s.d. obtained from three experiments performed in duplicate.
c. Cell viability assay showing the effect of FK866 on wild-type, Cas9+ (parental) and polyclonal G383del HDR-edited HAP1 cells after 72hr. Data points are normalized to untreated cells and represent averages ± s.d. obtained from three experiments performed in triplicate.
d. Knockout of *NAMPT* using CRISPR/Cas9-mediated genome editing shows *NAMPT* is important for cell survival. HAP1 or K-562 cells stably expressing SpCas9 were transfected with sgRNA plasmid pools targeting the indicated gene. The following day, cells were treated for 24hr with puromycin to select for transfected cells. Four days after selection, surviving cells were counted using trypan blue exclusion. *KIF11* is an essential gene and served as positive control, while non-targeting sgRNAs were used as negative control. Values represent means ± s.d. obtained from two independent experiments.

a

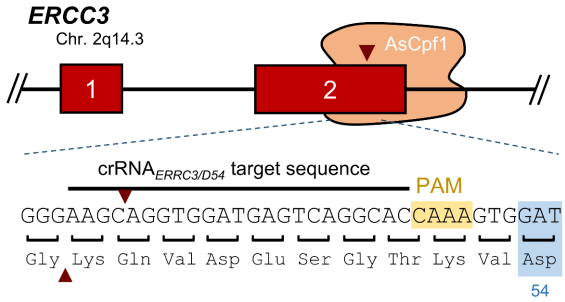


b

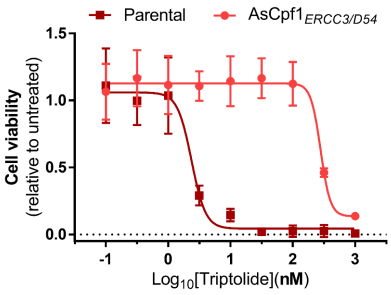
	Human NAMPT + KPT-9274
Data collection	
Space group	$P2_1$
Cell dimensions	
a, b, c (Å)	62.14, 109.17, 84.97
α, β, γ (°)	90.00, 96.63, 90.00
Resolution (Å)	50.00 – 2.05 (2.12 – 2.05)
R_{merge}	0.139 (0.557)
I / σI	9.0 (2.1)
Completeness (%)	100.0 (99.8)
Redundancy	3.8 (3.5)
Refinement	
Resolution (Å)	42.20 – 2.05
No. reflections	67,850
R_{work} / R_{free}	0.235 / 0.283
No. atoms	
Protein	7,478
Ligand/ion	184
Water	809
B-factors	
Protein	25.04
Ligand/ion	39.27
Water	31.15
Ramachandran Plot	
Favored (%)	96.75
Allowed (%)	3.25
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.66

Supplementary Figure 11: *details of the cocrystallization of KPT-9274 with NAMPT dimer.*
a. The electron density map of the ligand (yellow) in the NAMPT binding site at 1σ shows a perfect match between the ligand and the observed density.
b. Data collection and refinement statistics (molecular replacement).

a



b

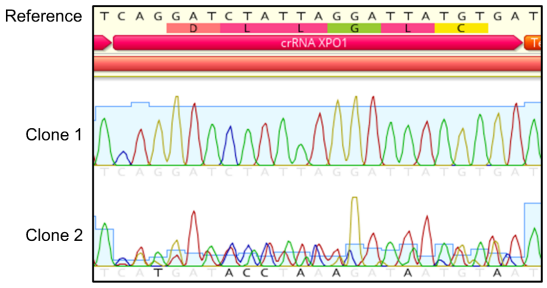


c

Type	Amino Acid Sequence	Read Frequency	CrispRVariants Identifier
Wild-type	…AAGKQVDES…	0.0%	no variant
Deletion	.....-...	12.8%	2:3D
Deletion	....-L...	8.6%	1:3D
Deletion	....-....	7.8%	-3:3D
Deletion	....--H..	4.6%	1:6D
Deletion	..-----...	3.7%	-8:12D
Frameshift	.....-MSQ	6.6%	2:4D
Frameshift	....--MSQ	4.0%	-2:7D
Other	-	51.9%	Other

Supplementary Figure 12: *AsCpf1 editing of the ERCC3^{D54} locus confers triptolide resistance.*
a. Overview of AsCpf1-mediated mutagenesis near the *ERCC3*^{D54} locus. D54, known to provide resistance upon mutation, is highlighted. The crRNA cutting site is denoted by red arrows.
b. Cell viability after 72hr of wild-type parental and polyclonal mutagenized HAP1 cells treated with different concentrations of triptolide. Signals are plotted relative to untreated control and points represent means ± s.d. obtained from three experiments performed in triplicate.
c. Amino-acid variants identified in the triptolide resistant HAP1 cells mutagenized at the *ERCC^{D54}* locus. Only reads with a read frequency above 3% are shown. Variants were obtained by CrispRVariants analysis of next-generation sequencing reads. Amino acids A41 to S49 are shown and no mutations were detected at D54.

a



b

crRNA sequence	Cpf1 cut site	Frequency
CAGGATCTATTAGGATTATGTGA	XPO1^{C528}	37.8%
CAGGATCTATTAGGATTATGTGA AGTAAAACCAACACAGAACAGAC	XPO1^{C528} RPS3A^{F138}	20.0%
CAGGATCTATTAGGATTATGTGA CCAGCATTCCTTGCTATTCCACC	XPO1^{C528} XPO1^{P867}	20.0%
CAGGATCTATTAGGATTATGTGA CTTACCTTCGCACTGGTTCCTTG	XPO1^{C528} XPO1^{R96}	17.8%
CAGGATCTATTAGGATTATGTGA TTCTTTATTTTCAGTAATCTATG	XPO1^{C528} XPO1^{M1}	2.2%
CAGGATCTATTAGGATTATGTGA CTTTCTTAGAGATCTTGCTGCCC	XPO1^{C528} ABL1^{A385}	2.2%

Supplementary Figure 13: *sequencing results of single-cell derived selinexor resistant colonies.*
a. Sequencing chromatogram of crRNAs present in single-cell derived selinexor resistant HAP1 colonies as obtained by Sanger sequencing. Each chromatogram denotes a different clone. The reference sequence above contains the *XPO1*^{C528} codon targeting crRNA sequence. The first clone contains a single crRNA targeting codon *XPO1*^{C528}. The second clone contains 2 crRNAs, one targeting codon *XPO1*^{C528} and one targeting codon *RPS3A*^{F138} (AGTAAAACCAACACAGAACAGAC). The clones were obtained from the pool of AsCpf1-mutageneized and selinexor resistant cells described in **Fig. 6d-h**.
b. All single-cell derived clones contained the *XPO1*^{C528} targeting crRNA and some clones were transduced with 2 crRNAs. Some mismatches in the AsCpf1 guide sequences were detected and are underlined.

Supplementary files available online

These files are available online with the manuscript:
Nature Communications 9, article number 502, 5 February (2018)
doi:10.1038/s41467-017-02349-8

https://www.nature.com/articles/s41467-017-02349-8

Supplementary Data 1
List of the 2,209 sgRNAs present in the SpCas9 tiling library covering 9 genes.

Supplementary Data 2
List of read counts of the sgRNAs present in HAP1 Cas9$^+$ cells transduced with the 9 gene SpCas9 tiling library before and after treatment with 8 nM Ispinesib.

Supplementary Data 3
List of the 40,517 sgRNAs present in the two SpCas9 FDA targets tiling sublibraries A and B, together covering 115 target genes.

Supplementary Data 4
List of the 24,435 sgRNAs present in the two SpCas9 Non-FDA targets tiling sublibraries C and D, together covering 75 target genes.

Supplementary Data 5
List of read counts of the sgRNAs present in HAP1 Cas9$^+$ cells transduced with the FDA tiling sublibrary B for SpCas9 before and after treatment with 30 nM Bortezomib.

Supplementary Data 6
List of read counts of the sgRNAs present in HAP1 and K-562 Cas9$^+$ cells transduced with the non-FDA tiling sublibraries C and D before and after treatment with 300 (HAP1) or 500 (K-562) nM KPT-9274.

Supplementary Data 7
Uncovered NAMPT mutations in transduced K-562 and HAP1 cells after treatment with KPT-9274.

Supplementary Data 8
List of the 1,100 crRNAs present in the AsCpf1 tiling library covering 10 genes.

Supplementary Data 9
List of read counts of the sgRNAs present in HAP1 Cas9$^+$ cells transduced with the AsCpf1 tiling library before and after treatment with 2 μM selinexor (KPT-330).

Supplementary Data 10
Overview of the primers, HDR templates and crRNA/sgRNA sequences used throughout the manuscript.

Supplementary Data 11
R scripts used for data analysis and library generation.

CHAPTER V

General discussion

Background

Analysis of the pharmaceutical industry and their output over the last few decades indicates that new drug candidates often fail during late-stage development due to a lack of efficacy[1, 31]. Major reasons for this failing have been attributed to wrong assumptions on target to disease phenotype relations and to a lack of robust target deconvolution and validation to understand a drug's mechanism of action[1, 36]. Due to this lack of understanding and the recent resurgence of phenotypic screening in drug discovery, the requirement of characterizing and understanding drug-target interactions to support the drug development process has become stronger than ever. Not surprisingly, many methods have already been developed to validate and identify drug-target interactions[37-39, 44]. These methods can be divided into three major groups: biochemical methods, computational methods and genetic methods, each with their own benefits and pitfalls. Chemical-genetic relations can provide strong evidence for target validation as they examine drug-target interactions in a living cellular context. Therefore, the isolation and validation of drug-resistance mutations in a cellular context is regarded as the gold standard proof for target confirmation[38, 200]. While this process is widely used in prokaryotes, lower eukaryotes and viruses, its application in mammalian cells has been limited due to the complexity of the large multiploid genome. However, after pioneering work in yeast, chemical-genetic methods have now started to come to fruition in mammalian cells with the advent of next-generation sequencing technologies, RNA interference and, more recently, CRISPR/Cas-mediated genome editing[37, 38].

The discovery that short RNA fragments can potently suppress gene transcripts by RNA interference (RNAi) allowed researchers for the first time to study chemical-genetic interactions on a large scale in mammalian cells[160, 161]. Although RNAi provides a powerful and controllable methodology, it is notorious for its noisy data attributed to off-target effects of the RNAi reagents themselves and poor knockdown efficiencies of individual target genes[165]. These shortcomings have in part been addressed over time by a better understanding of cellular RNA interference and by optimization of the RNAi reagents[163, 164]. In addition, new statistical tools have been developed to better deconvolute the results obtained with RNAi. However, the large size of improved RNAi reagent libraries and the difficulty to validate and deconvolute the results can still prohibit RNAi's utility for performing genome-wide screens.

The arrival of CRISPR/Cas-mediated genome editing has provided researchers with an easy to use toolbox to manipulate the genetic information in a living cell at its source with unprecedented resolution and precision[214, 423, 458]. CRISPR/Cas genetic engineering utilizes a bacterial endonuclease (e.g. SpCas9) that is guided to its target DNA by an engineered small single-stranded guide RNA (sgRNA) that is complementary to the targeted DNA sequence. When the endonuclease has bound its target, it introduces a DNA double strand break, which triggers the cell to repair the DNA damage through non-homologous end-joining (NHEJ) or homology-directed repair (HDR)[247]. NHEJ repair of CRISPR-induced breaks is error-prone and continues until mutations have been introduced that abolish targeting of the endonuclease to the DNA locus due to a loss of sequence complementary between the sgRNA and DNA. In protein coding sequences, these mutations often result in frameshift mutations, knocking out the function of the encoded protein. HDR, on the other hand, is much more precise but less efficient and uses a template DNA similar to the damaged DNA to repair the damage. In addition to the active endonuclease, different variants of CRISPR have been developed that rely on inactivated mutants of the endonuclease (e.g. dead SpCas9; dCas9)[213]. These variants essentially function as a programmable DNA-binding protein that can be used to recruit other functional protein domains to DNA target loci without cutting of the DNA. The ever-expanding CRISPR toolbox thus has the potential to systematically control genetic information and the concepts developed from chemical-genetic efforts in yeast and with RNA interference are now being applied to CRISPR/Cas-mediated genome editing. In this thesis, we investigated the use of CRISPR/Cas-mediated genome editing for the validation and identification of a drug's cellular target.

Validation of the cellular drug-target selectivity of XPO1 inhibitors

It is important to validate that the observed bioactivity of a candidate drug is caused by interaction of the drug with its suggested target to understand the drug's mechanism of action. In chapter III of this thesis, we investigated whether we can use CRISPR/Cas-mediated genome editing to validate drug-target interactions in a human cellular context. For this purpose, we turned to the candidate drug selinexor (KPT-330), which is the clinical candidate of a promising class of anticancer drugs targeting the essential exportin-1 (XPO1)[359, 399-409]. Amongst the human exportins, XPO1 facilitates the active transport of by far the largest number of proteins from the nucleus to the cytoplasm[340-344, 410, 411]. Its cargo proteins include transcription factors, tumor suppressors and cell cycle regulators such as p53, IκB, pRb, APC and FOXO family proteins. XPO1 therefore controls many vital aspects of cellular physiology and has garnered interest as a potential anticancer target[331, 410, 411]. Selinexor has been shown to interact covalently with the cysteine$_{528}$ residue located in the hydrophobic NES-binding pocket of XPO1[399, 404]. This NES-binding pocket is essential for XPO1 function as it is required for the recognition of its cargo proteins through their nuclear export signal sequence (NES). However, the causality between inhibition of XPO1 and the observed anticancer mechanism of action of selinexor has not been demonstrated in human cells. To study the selinexor-XPO1 drug target selectivity, we applied CRISPR/SpCas9-induced homology directed repair to mutate the cysteine$_{528}$ residue of endogenous XPO1 in leukemia cell lines sensitive to selinexor to generate isogenic XPO1^{C528S} mutant cell lines. Although this cysteine residue is highly conserved across higher eukaryotes, we opted to mutate this cysteine to serine, as a serine residue at this location in the NES-binding cleft has been shown to be viable in some yeast species[353, 373, 387]. We were able to generate both hetero- and homozygous XPO1 mutants that were highly resistant to treatment with selinexor. In more detail, the serine substitution of cysteine$_{528}$ in XPO1 conferred cells with strong resistance to the typical effects induced by treatment with selinexor. Mutant cells were resistant to apoptosis induction, cell cycle arrest, proteasomal degradation of XPO1, inhibition of XPO1-mediated nuclear export and direct binding of XPO1 by selinexor. In addition, heterozygous mutation of the cysteine residue was sufficient to cause selinexor resistance, confirming a dominant resistance phenotype of the C528S mutation. Together, these results demonstrate and validate that the observed anticancer activity of selinexor is specific for modulating XPO1 function by interacting with the cysteine$_{528}$ residue of XPO1. Moreover, they highlight the potential of using CRISPR/Cas-mediated genome editing for validation of drug-target interactions by altering the endogenous genetic information at its source. As the CRISPR/Cas9 endonuclease can easily be reprogrammed, the approach should be applicable to almost any region in the genome. Finally, we obtained both hetero- and homozygous mutants, showing the approach also allows to study both dominant and recessive resistance mutations.

Clinical development of XPO1-mediated nuclear export inhibitors

The improved small molecule *N*-azolylacrylates that inhibit XPO1-mediated nuclear export (SINE) show potent anticancer activity *in vitro* and *in vivo* and to date >2,000 patients have been treated with selinexor in Phase 1,2 and 3 clinical trials[390]. These clinical trials focused on the treatment of a broad range of hematological malignancies and solid tumors, including acute myeloid leukemia (AML), pancreatic cancer, multiple myeloma (MM), diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma (MCL), myelodysplastic syndrome (MDS), acute lymphoblastic leukemia (ALL), ovarian and cervical carcinoma, lung cancer and glioblastoma. At this moment, the focus lies at clinical trials focused on MM, DLBCL, liposarcoma, and glioblastoma (https://www.karyopharm.com/pipeline/). During the clinical trials it was observed that selinexor induces constitutional symptoms such as weight loss, fatigue, anorexia, confusion and thrombocytopenia[389, 408, 409, 419, 459, 460]. However, tolerability to selinexor has improved by supportive care with appetite stimulants, antinausea agents, and by limiting the dosing to a schedule of twice/thrice weekly.

With eye on improving the tolerability of the SINE, second generation SINEs have been developed. The main candidate of this second generation of inhibitors is eltanexor (KPT-8602) and subsequent studies, including one by our group, have shown that eltanexor is just as, or even more effective as, selinexor in killing leukemia cells *in vitro* and in mouse models of cancer[459, 461, 462]. Moreover, we were able to use our XPO1^{C528S} cell lines to show eltanexor is just as selective for the cysteine$_{528}$ residue in XPO1 with our mutant cells being over 100-fold resistant[461]. Eltanexor was designed to bind more reversibly to cysteine$_{528}$ than selinexor, but this could not be demonstrated[461, 463]. Instead, it showed strongly improved tolerability profiles in mice that have been attributed to the drug's inability to cross the blood-brain barrier[459, 462]. This allows for much more frequent dosing of eltanexor, possibly up to daily dosing regimens. As a result, eltanexor has now moved to phase 1 clinical trials and will be evaluated for its potential in multiple myeloma and hematological malignancies. Finally, besides the SINE, the SL-801 (CBS9106 CanBas Co. Ltd.) covalent XPO1 inhibitor has also entered phase 1 clinical trials targeted against cancer, although it remains to be seen how this compound will develop clinically.

Searching for XPO1 inhibitors that do not target the cysteine$_{528}$ residue

Although selinexor and eltanexor might hold great potential as future anticancer therapeutics, it must be noted that they still bind covalently to the cysteine$_{528}$ residue in XPO1. Historically, covalent drugs have been attributed with an increased toxicity and higher occurrence of off-target effects and adverse events[464, 465].Although we demonstrated XPO1 inhibitors do not kill cancer cell lines carrying a cysteine$_{528}$ mutation in XPO1 *in vitro*, we cannot exclude the covalent nature of XPO1 inhibitors might cause side-effects in a complete organism. As a result, it might be beneficial to discover non-covalent XPO1 inhibitors to prevent the observed toxicity associated with current XPO1 inhibitors. In addition, the cysteine$_{528}$ residue is present at the center of the NES-binding cleft of XPO1. Thus, binding to this residue by current XPO1 inhibitors is thought to inhibit the XPO1-mediated nuclear transport of a large variety of cargo proteins. More selective XPO1 inhibitors inhibiting only a select subset of cargo proteins might be able to induce potent anticancer effects without causing side effects. Interestingly, structural regions surrounding the NES-binding cleft of XPO1 are used by snurportin-1 and maybe several other cargo proteins to interact with XPO1[374]. These interactions involve binding to HEAT repeats 11-17, which display distinct charged and conserved patches on the outer surface of XPO1 near the NES-binding cleft. These patches allow the binding of cargo proteins by opposing surface charge and could, in addition to binding within the NES-binding cleft, play an important role in the binding of some cargo proteins to XPO1[374]. However, the list of XPO1 cargo proteins is still incomplete and the interactions of many full-length cargo proteins with XPO1 are very poorly documented. Therefore, obtaining an extensive picture of XPO1-cargo interactions will require additional crystallographic analysis and the study of many full-length cargo proteins. Nevertheless, the pockets surrounding the NES-binding cleft might provide additional anchor points for the discovery or rational design of new XPO1 inhibitors. The XPO1^{C528S} mutant cell lines reported here could provide a strong basis for drug discovery efforts targeted at the discovery of such XPO1 inhibitors. Indeed, they do not contain the cysteine$_{528}$ residue, preventing the discovery of additional covalent binders of this cysteine residue.

XPO1 also plays a role during mitosis[466, 467]. Together with nucleoporins, survivin and RanGTP, XPO1 facilitates the recruitment of nucleation factors to the kinetochores of mitotic chromosomes and is required for the correct formation of microtubule networks during spindle formation. During mitosis, XPO1 can be phosphorylated by the mitotic kinase CDK1/cyclin B to enhance its ability in targeting RanBP2 and RanGAP1 to the mitotic spindle[468]. In addition, XPO1 functions as a transporter and chaperone that enables the recruitment of BRCA1/BARD1 and other proteins to the centrosomes to facilitate centrosome duplication by recognition of their nuclear export signals[469]. Inhibition of XPO1 function with leptomycin B during early

metaphase leads to formation of acentriolar spindle poles that disrupt the formation of microtubule networks, although centrosome duplication is not affected[470]. Besides the role of XPO1 during mitosis, XPO1 has also been implicated in ribosome biogenesis and can localize to the nucleolus[471, 472]. It has been associated with the activity of RNA polymerase I at the nucleoli and binds to CPEB1 and NMD3, which contain nuclear export signals and can colocalize with XPO1 at the nucleolus[473, 474]. Moreover, many ribosomal proteins contain a nuclear export signal and although these sequences are essential for their export to the cytoplasm, they have been implicated to target XPO1 to the nucleoli. The exact role of XPO1 at the nucleoli remains elusive, but inhibition of XPO1 by small molecule inhibitors of nuclear export disrupts ribosomal biogenesis and induces ribosomal stress in multiple myeloma and mantle cell lymphoma[475, 476]. Although the anticancer activity of XPO1 inhibitors is mainly thought to occur through relocalization of important growth regulatory and tumor suppressor factors to the nucleus during interphase, it would be possible the potent bioactivity of XPO1 inhibitors in cancer cells is in part explained by interfering with the mitotic and ribosomal-related functions of XPO1. However, not much is known about the effect of XPO1 inhibitors on these alternative functions, especially in a clinical setting. Thus, it would be interesting to examine to what extent modulation of these functions contributes to the observed bioactivity of the XPO1 inhibitors.

CRISPR/Cas mutagenesis scanning for target deconvolution

Genetic screens hold great potential for target deconvolution as they allow to examine drug-target interactions in a cellular context[38]. Different approaches to chemical-genetic screens exist and rely on different principles. As the biological activity of a candidate drug is correlated with the extent of interaction of the drug with its target, decreasing or increasing the abundance of the target should modulate the drug response. Based on this principle, genetic overexpression screens aim to identify genes whose increased concentration or expression modulates the activity of a drug[193, 194]. However, the potential of these screens is often hampered by the identification of indirect mechanisms of resistance as increased expression of drug efflux pumps and general pro-survival and growth regulatory genes can suppress the activity of a drug, although such proteins are often not the drug's direct target. Besides these overexpression screens, it is also possible to reduce the expression of genes using RNA interference or gene-knockout screening to identify genes that modulate the drug response[164, 214, 222]. However, as with overexpression screens, these genetic screens often fail to identify the direct cellular target of the drug and typically identify modulatory genes that act downstream of the drug's target. In addition, gene knockout screening cannot be used to probe essential genes as a complete removal of these genes is lethal to the cell. The gold standard proof for a drug's cellular target is the isolation and identification of mutations that confer drug resistance in a cellular context[200]. As such, other genetic methods aim to identify resistance mutations in the target protein of a drug. Typically, these methods rely on applying potent mutagens or on cell lines that are inherently mutagenic due to defects in their DNA polymerases or repair proteins to generate multiple random mutations across the genome. Common mutagens comprise DNA alkylating agents such as *N*-ethyl-*N*-nitrosourea (ENU) and ethyl methane sulfonate (EMS) or the use of UV irradiation[38, 204, 205]. However, due to the large size and complexity of the mammalian genome, the deconvolution of the relevant resistance mutations generated by these approaches is costly and troublesome, requiring whole exome sequencing of multiple independent clones[204, 205]. Moreover, as obtaining a similar mutation in multiple alleles of the same gene is challenging, these methods are restricted to uncover dominant resistance mutations in multiploid cells and are therefore mainly used in haploid systems. These examples make it clear that the field requires new approaches for drug target deconvolution that can simplify the identification of drug-target interactions.

In chapter IV of this thesis we established a new genetic screening approach to explore drug-target interactions of anticancer drugs. We reasoned the introduction of DSBs by targeted CRISPR/Cas endonucleases and the subsequent error-prone repair via non-homologous end-joining (NHEJ) could be exploited for mutagenesis of essential genes to facilitate drug resistance selection. If feasible, the sgRNA sequences used to guide the CRISPR endonuclease can be used as a marker to identify the target region of mutagenesis. This would allow to circumvent whole exome sequencing of individual clones to deconvolute relevant drug resistance mutations. Thus, we set out to validate this principle by applying CRISPR/Cas-induced mutagenesis to obtain drug resistance against 3 well-characterized anticancer drugs. Using these established drugs, we showed CRISPR/Cas-induced NHEJ repair forms a plethora of mutations that can be used to select for functional in-frame drug resistance mutations in a drug's target gene. We then established this simple concept can be applied as a genetic screening strategy using large sgRNA tiling libraries spanning the entire coding sequence of multiple essential genes to identify the cellular target protein of a drug. We correctly identified PSMB5 as the target protein of the well-known anticancer agent bortezomib. In addition, we were able to identify and validate NAMPT as the cellular target of a novel anticancer compound still in development. Importantly, the target sequence of the enriched sgRNAs quickly guided us to the target gene in both screens. Our results demonstrate CRISPR/Cas-induced NHEJ repair generates gain-of-function mutations and show this concept can be applied in a rapid and powerful genetic screen to obtain gold standard proof for drug-target interactions in a cellular context.

Considerations and improvements for CRISPR/Cas mutagenesis scanning

Although the initial success of CRISPRres is promising, the approach may be improved. First, in its current format we were able to scan 75 genes in a single screen. If the target gene of the drug of interest is not present in the list of 75 genes, the screen will fail or lead to the identification of false positive sgRNA hits if drug resistant cells are obtained. False positives can also be discovered if the initial population of cells already contained drug resistant cells. Therefore, it is important to repeat the genetic screen to filter out for passenger sgRNAs that are not related to the observed drug resistance. In addition, the enriched sgRNA hits should be validated independently using a simple transfection experiment. Nevertheless, the human genome contains ~20,000 genes of which 8,000-10,000 are expressed at a given moment. It would be time consuming to scan these using pools of 75 target genes. However, we suggest our method should initially only be applied to essential genes. This suggestion assumes that if a drug's bioactivity is caused by interacting with a non-essential target protein, this target protein is very likely to be identified by a genetic loss-of-function screen. Cells express a much lower number of essential genes, with estimates ranging between 1,500~2,500[269]. Although this is still a relatively large list, we envisage the screening method can be applied on a list of candidate genes identified after a first round of target deconvolution by e.g. affinity chemical proteomics, overexpression screens, or comparative studies such as CMap[36, 38, 44, 449]. The approach could also be applied to a set of predefined genes to only select for compounds that modulate the function of pathways or genes of interest. Moreover, there is room to improve the methodology further. These improvements might make it possible to scan a few hundred genes in a single experiment. Of course, it remains to be seen how such improved systems will perform.

The general performance of the approach relies on the efficient generation of gain-of-function in-frame indels or base substitutions during NHEJ repair[214]. While this is sufficiently effective for the drugs and sites investigated here, a range of different strategies may be examined to increase the occurrence of in-frame indels or base substitutions. For example, small molecules modulating the cellular DNA repair mechanism[477, 478], appending sgRNA's with a hairpin aptamer to recruit additional effector proteins[289], effector domain insertions into the Cas9 endonuclease structure[479] or the use of Cas endonuclease orthologs[458, 480] may be applied to guide the NHEJ repair of the double strand break towards increased in-frame indels or substitutions. If the occurrence of in-frame mutations can be increased, the number of cells needed for an experiment will drop. This provides the opportunity to expand the list of target genes in the library while keeping the total number of required cells the same.

Currently, the methodology utilizes stable cell lines continuously expressing the CRISPR/Cas endonuclease and a constitutive active U6 promotor for the sgRNAs. The moment these cells are transduced with a lentiviral particle containing an sgRNA, cells will almost immediately start expressing the sgRNA. This sgRNA is immediately used by the already present Cas endonuclease to cut and mutagenize the DNA, even before the cells can divide. With the current setup, transduced cells are grown for a period of 4-5 days under selection to remove untransduced cells before they are challenged with the candidate drug. Therefore, at the time of drug treatment, the population will contain many cells with the same mutation, as the originally mutagenized cells have now divided multiple times. In other words, the genetic variation present in the population at the time of drug treatment is relatively low. Now, if cells are granted some time during the incubation period to divide before the CRISPR endonuclease or sgRNA are expressed, the number of cells containing a unique mutation at the time of drug challenge can only increase. Thus, another way to increase the power of the methodology would be to use an inducible CRISPR/Cas endonuclease or sgRNA system. Nevertheless, it would also be possible to upscale the experiment by increasing the number of starting cells at the start of the experiment to increase the genetic variation in the population. However, this will require a more advanced and extensive practical workflow, which would benefit greatly from optimized cell culture facilities using for example automated robotics or small-scale bioreactors.

The resolution and potential of the methodology are further defined by the distribution of the protospacer adjacent motif associated with the used CRISPR/Cas endonuclease across the genome. On average, we observed that 71% of the coding sequence of our target genes falls within a 9 bp window around SpCas9 cut sites. This means ~30% of the coding sequences is not covered. If a drug binds to residues located in these unrepresented regions, the method will likely be unable to identify the target. Fortunately, we have shown the methodology is compatible with the CRISPR/AsCpf1 endonuclease, which recognizes the TTTV PAM[226]. Moreover, as the genetic variation is generated by the cellular NHEJ repair machinery, the approach is likely to be compatible with other CRISPR/Cas endonucleases as long as they induce DNA double strand breaks. Thus, a combination of CRISPR/Cas endonucleases recognizing different PAMs might enable a close to complete coverage of coding sequences.

Finally, CRISPRres is unlikely to uncover the target of a drug if the bioactivity of this drug is caused by modulation of multiple target genes. The method is set up to facilitate transduction of a single cell with only 1 sgRNA. This allows transduced cells to generate mutations at only a single site, preventing them from generating resistance mutations in multiple genes. Even when multiple sgRNAs enter the same cell, it is unlikely this will lead to drug resistance as the chance that this cell contains sgRNAs that target different target proteins and at the same time lead to functional resistance mutations will be (very) low when using large multi-gene tiling libraries. Importantly, the system also utilizes active CRISPR/Cas endonucleases. Thus, when multiple sgRNAs enter the same cell, a DNA double strand break will be introduced at multiple sites. This facilitates formation of chromosomal rearrangements and translocations[481]. Furthermore, if multiple sgRNAs target the same chromosome, large deletions are likely to be generated because large parts of the DNA will be excised from the chromosome[433]. These genetic byproducts are unlikely to contribute to drug resistance, in fact they are more likely to negatively impact cell viability. Nevertheless, pairs of sgRNAs targeting different sites have been used to introduce double knockouts[482, 483]. As such, performing small scale CRISPRres combinatorial screens might be feasible, although this remains to be investigated.

CRISPR/Cas mutagenesis scanning for protein evolution in a native context

CRISPR/Cas mutagenesis scanning could also be applied to engineer proteins with new functions as the main concept underlying the methodology is the generation of a diverse collection of variants at the CRISPR/Cas target site. Current protein evolution approaches rely on the generation of diverse populations using either global mutagenesis with radiation and chemically induced DNA damage or plasmid libraries coupled to replication with low-fidelity polymerases or phage-displays[484, 485]. However, these approaches are not feasible in a mammalian context as mammalian cells replicate slowly and do not propagate plasmids. In addition, the large size of mammalian genomes hinders the utility of global mutagenesis as they generate semi-random mutations across the genome. Moreover, when mammalian proteins are engineered in bacteria, yeast or bacteriophages, they are often not functional anymore in the native host. Hence, obtaining functional protein variants in a mammalian cellular context is difficult. Recent work demonstrated CRISPR/Cas9-induced homologous recombination can be applied to generate protein variants in mammalian cells using oligonucleotide libraries containing diverse sets of mutations[486-488]. However, this strategy requires synthesis of a mutant donor library for each targeted site and relies on HDR, an inefficient repair process. CRISPR/Cas mutagenesis scanning would allow for rapid protein evolution in a native context. However, its potential in its current format would be limited as the CRISPR cut site and DNA recognition site are coupled. Once mutations have been selected out, the system will be unable to recut the mutated site as the sgRNA is not complementary to the target site anymore. Thus, to generate additional mutations in the target protein, a new round of evolution is required. Moreover, as the system utilizes DNA DSBs, simultaneous mutagenesis of multiple sites in the same protein will be challenging. Nevertheless, with optimization it might prove a valuable tool.

Conclusion

Uncovering the cellular target and the underlying mechanism of action of a candidate drug remains an important aspect of the drug discovery and development process. Target knowledge helps predict tractability and validity of the candidate drug in a clinical setting. It allows for improved hit-to-lead optimization and supports biomarker development. However, deconvolution of the cellular target of a candidate drug remains a major challenge. Various methods have been developed over time and although powerful, none of them provides a general solution. Dissecting the functional genetic requirements for bioactivity of a candidate drug remains one of the most powerful pieces of proof to validate and identify drug-target interactions. However, especially in the context of mammalian cellular systems, chemical-genetic methods have lagged behind. Fortunately, chemical-genetic screening has finally started to pick up steam in a mammalian context with the advent of next-generation sequencing technologies, RNA interference and, more recently, CRISPR/Cas-mediated genome editing.

In this thesis we explored the use of CRISPR/Cas-mediated genome editing for the validation and identification of drug-target interactions in a cellular context. Using CRISPR/Cas-induced homology directed repair to mutate the cysteine$_{528}$ residue of XPO1, we were able to confirm the drug-target selectivity of clinical XPO1 inhibitors. We also discovered that endogenous substitution of the cysteine$_{528}$ residue affects cellular growth and XPO1 function, suggesting this residue might be important for the function of human XPO1. However, further studies are required to better understand the role of this cysteine residue in XPO1's complex biology. We then established CRISPRres, a new method for drug target deconvolution based on the gold standard proof for target deconvolution and CRISPR/Cas-induced non-homologous end joining. We show that CRISPR/Cas-induced NHEJ leads to the generation of functional in-frame variants which can be used to derive endogenous mutant protein variants that are resistant to inhibition by a small molecule drug, even when these proteins are essential for cell survival. We then applied this concept as a genetic screening approach to probe entire coding sequences of a list of genes to identify NAMPT as the cellular target of an investigational anticancer compound. We also showed this genetic approach is compatible with both the SpCas9 and AsCpf1 CRISPR endonucleases. This is important as the protospacer adjacent motif recognized by the used endonuclease dictates the resolution of the approach. Although the methodology has already shown its usefulness, further optimization is required to increase its power and resolution. Overall, our results show CRISPR/Cas genome editing provides a strong toolbox for validation and identification of drug-target interactions in a cellular context.

Finally, our study and numerous other examples in the literature show how diverse and developed CRISPR/Cas genetic screening has already become over a timespan of just a few years. Yet, there are no clear signs of a slowed down development and the CRISPR/Cas toolbox will no doubt be expanded even further. As the technology continues to develop, we can expect to obtain an almost unlimited control over the genome in almost any organism. Especially in a mammalian context, CRISPR/Cas has revolutionized genetic screening and will likely replace or at least provide very strong alternatives to previous methods (**Table 1**), besides continuing to provide new functional-genetic approaches to study almost all aspects of biology.

Current genetic methods	CRISPR/Cas alternative
RNA interference *(Lentiviral shRNA)*	**CRISPR interference** *(Lentiviral sgRNA/dCas9-KRAB)*
Loss of function screening in haploid cells *(Transposon, lentiviral gene traps)*	**CRISPR knockout screening** *(Lentiviral sgRNA/Cas9)*
Overexpression screening *(Transposon promotors, lentiviral ORFs)*	**CRISPR activation** *(Lentiviral sgRNA/dCas9-VPR/SAM)*
Point mutations *(ENU/EMS, UV)*	**CRISPRres, CRISPR-X, base-editors** *(Lentiviral sgRNA/dCas9-AID*Δ)*

Table 1: *overview of CRISPR/Cas alternatives for genetic screening methods in mammalian cells.*

Acknowledgements and contributions

Acknowledgements
Scientific acknowledgements are mentioned in the respective result chapters III and IV.

Conflict of interest statement
Conflict of interest statements are mentioned in the result chapters III and IV.

Personal contribution
Jasper Edgar Neggers has designed and written this manuscript. He also designed the figures presented in this manuscript. The author contributions for the results described in Chapter III and IV are described in detail in the respective chapters.

References

1. Plenge, R.M. Disciplined approach to drug discovery and early development. *Sci Transl Med* **8**, 349ps315 (2016).
2. Andrade, E.L. et al. Non-clinical studies required for new drug development - Part I: early in silico and in vitro studies, new target discovery and validation, proof of principles and robustness of animal studies. *Braz J Med Biol Res* **49**, e5644 (2016).
3. Hughes, J.P., Rees, S., Kalindjian, S.B. & Philpott, K.L. Principles of early drug discovery. *Br J Pharmacol* **162**, 1239-1249 (2011).
4. Paul, S.M. et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* **9**, 203-214 (2010).
5. DiMasi, J.A., Grabowski, H.G. & Hansen, R.W. Innovation in the pharmaceutical industry: new estimates of R&D costs. *J Health Econ* **47**, 20-33 (2016).
6. Roses, A.D. Pharmacogenetics in drug discovery and development: a translational perspective. *Nat Rev Drug Discov* **7**, 807-817 (2008).
7. Davis, A.M., Plowright, A.T. & Valeur, E. Directing evolution: the next revolution in drug discovery? *Nat Rev Drug Discov* **16**, 681-698 (2017).
8. Mayr, L.M. & Bojanic, D. Novel trends in high-throughput screening. *Curr Opin Pharmacol* **9**, 580-588 (2009).
9. Jensen, P.R. Natural products and the gene cluster revolution. *Trends Microbiol* **24**, 968-977 (2016).
10. Stone, M.J. & Williams, D.H. On the evolution of functional secondary metabolites (natural products). *Mol Microbiol* **6**, 29-34 (1992).
11. Lachance, H., Wetzel, S., Kumar, K. & Waldmann, H. Charting, navigating, and populating natural product chemical space for drug discovery. *J Med Chem* **55**, 5989-6001 (2012).
12. Franzini, R.M. & Randolph, C. Chemical space of DNA-encoded libraries. *J Med Chem* **59**, 6629-6644 (2016).
13. Lavecchia, A. & Di Giovanni, C. Virtual screening strategies in drug discovery: a critical review. *Curr Med Chem* **20**, 2839-2860 (2013).
14. Plowright, A.T. et al. Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. *Drug Discov Today* **17**, 56-62 (2012).
15. Baxter, K. et al. An end to the myth: there is no drug development pipeline. *Sci Transl Med* **5**, 171cm171 (2013).
16. Wagner, J. et al. A dynamic map for learning, communicating, navigating and improving therapeutic development. *Nat Rev Drug Discov* **17**, 150 (2018).
17. Wagner, J.A. et al. Application of a dynamic map for learning, communicating, navigating, and improving therapeutic development. *Clin Transl Sci* **11**, 166-174 (2018).
18. Smietana, K., Siatkowski, M. & Moller, M. Trends in clinical success rates. *Nat Rev Drug Discov* **15**, 379-380 (2016).
19. Moffat, J.G., Vincent, F., Lee, J.A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. *Nat Rev Drug Discov* **16**, 531-543 (2017).
20. Horman, S.R., Hogan, C., Delos Reyes, K., Lo, F. & Antczak, C. Challenges and opportunities toward enabling phenotypic screening of complex and 3D cell models. *Future Med Chem* **7**, 513-525 (2015).
21. Moffat, J.G., Rudolph, J. & Bailey, D. Phenotypic screening in cancer drug discovery - past, present and future. *Nat Rev Drug Discov* **13**, 588-602 (2014).
22. Wagner, B.K. & Schreiber, S.L. The power of sophisticated phenotypic screening and modern mechanism-of-action methods. *Cell Chem Biol* **23**, 3-9 (2016).
23. Vincent, F. et al. Developing predictive assays: the phenotypic screening "rule of 3". *Sci Transl Med* **7**, 293ps215 (2015).
24. Swinney, D.C. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin Pharmacol Ther* **93**, 299-301 (2013).
25. Wishart, D.S. Improving early drug discovery through ADME modelling: an overview. *Drugs R D* **8**, 349-362 (2007).

26. Kassel, D.B. Applications of high-throughput ADME in drug discovery. *Curr Opin Chem Biol* **8**, 339-345 (2004).
27. Umscheid, C.A., Margolis, D.J. & Grossman, C.E. Key concepts of clinical trials: a narrative review. *Postgrad Med* **123**, 194-204 (2011).
28. Martin, L., Hutchens, M., Hawkins, C. & Radnov, A. How much do clinical trials cost? *Nat Rev Drug Discov* **16**, 381-382 (2017).
29. Hay, M., Thomas, D.W., Craighead, J.L., Economides, C. & Rosenthal, J. Clinical development success rates for investigational drugs. *Nat Biotechnol* **32**, 40-51 (2014).
30. Shih, H.P., Zhang, X. & Aronov, A.M. Drug discovery effectiveness from the standpoint of therapeutic mechanisms and indications. *Nat Rev Drug Discov* **17**, 19-33 (2018).
31. Waring, M.J. et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov* **14**, 475-486 (2015).
32. Harrison, R.K. Phase II and phase III failures: 2013-2015. *Nat Rev Drug Discov* **15**, 817-818 (2016).
33. Pregelj, L., Verreynne, M.L. & Hine, D. Changes in clinical trial length. *Nat Rev Drug Discov* **14**, 307-308 (2015).
34. Lang, T. & Siribaddana, S. Clinical trials have gone global: is this a good thing? *PLoS Med* **9**, e1001228 (2012).
35. Vauquelin, G. & Charlton, S.J. Long-lasting target binding and rebinding as mechanisms to prolong in vivo drug action. *Br J Pharmacol* **161**, 488-508 (2010).
36. Bunnage, M.E., Gilbert, A.M., Jones, L.H. & Hett, E.C. Know your target, know your molecule. *Nat Chem Biol* **11**, 368-372 (2015).
37. Jost, M. & Weissman, J.S. CRISPR approaches to small molecule target identification. *ACS Chem Biol* **13**, 366-375 (2018).
38. Nijman, S.M. Functional genomics to uncover drug mechanism of action. *Nat Chem Biol* **11**, 942-948 (2015).
39. Schenone, M., Dancik, V., Wagner, B.K. & Clemons, P.A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* **9**, 232-240 (2013).
40. Bunnage, M.E. Getting pharmaceutical R&D back on target. *Nat Chem Biol* **7**, 335-339 (2011).
41. Blomme, E.A. & Will, Y. Toxicology strategies for drug discovery: present and future. *Chem Res Toxicol* **29**, 473-504 (2016).
42. Bowes, J. et al. Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nat Rev Drug Discov* **11**, 909-922 (2012).
43. Swinney, D.C. & Anthony, J. How were new medicines discovered? *Nat Rev Drug Discov* **10**, 507-519 (2011).
44. Terstappen, G.C., Schlupen, C., Raggiaschi, R. & Gaviraghi, G. Target deconvolution strategies in drug discovery. *Nat Rev Drug Discov* **6**, 891-903 (2007).
45. Schirle, M. & Jenkins, J.L. Identifying compound efficacy targets in phenotypic drug discovery. *Drug Discov Today* **21**, 82-89 (2016).
46. Chan, J.N., Nislow, C. & Emili, A. Recent advances and method development for drug target identification. *Trends Pharmacol Sci* **31**, 82-88 (2010).
47. Cong, F., Cheung, A.K. & Huang, S.M. Chemical genetics-based target identification in drug discovery. *Annu Rev Pharmacol Toxicol* **52**, 57-78 (2012).
48. Mallidi, S., Spring, B.Q. & Hasan, T. Optical imaging, photodynamic therapy and optically triggered combination treatments. *Cancer J* **21**, 194-205 (2015).
49. Puyo, S., Montaudon, D. & Pourquier, P. From old alkylating agents to new minor groove binders. *Crit Rev Oncol Hematol* **89**, 43-61 (2014).
50. Rix, U. & Superti-Furga, G. Target profiling of small molecules by chemical proteomics. *Nat Chem Biol* **5**, 616-624 (2009).
51. Pan, S., Zhang, H., Wang, C., Yao, S.C. & Yao, S.Q. Target identification of natural products and bioactive compounds using affinity-based probes. *Nat Prod Rep* **33**, 612-620 (2016).
52. Su, Y. et al. Target identification of biologically active small molecules via in situ methods. *Curr Opin Chem Biol* **17**, 768-775 (2013).
53. Lee, H. & Lee, J.W. Target identification for biologically active small molecules using chemical biology approaches. *Arch Pharm Res* **39**, 1193-1201 (2016).

54. Oda, Y. et al. Quantitative chemical proteomics for identifying candidate drug targets. *Anal Chem* **75**, 2159-2165 (2003).
55. Ong, S.E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**, 376-386 (2002).
56. Gygi, S.P. et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-999 (1999).
57. Bantscheff, M. et al. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol* **25**, 1035-1044 (2007).
58. Wang, D.Y. et al. Target identification of kinase inhibitor alisertib (MLN8237) by using DNA-programmed affinity labeling. *Chemistry* **23**, 10906-10914 (2017).
59. Chen, X. et al. Target identification with quantitative activity based protein profiling (ABPP). *Proteomics* **17** (2017).
60. Pichler, C.M., Krysiak, J. & Breinbauer, R. Target identification of covalently binding drugs by activity-based protein profiling (ABPP). *Bioorg Med Chem* **24**, 3291-3303 (2016).
61. Speers, A.E., Adam, G.C. & Cravatt, B.F. Activity-based protein profiling in vivo using a copper(i)-catalyzed azide-alkyne [3 + 2] cycloaddition. *J Am Chem Soc* **125**, 4686-4687 (2003).
62. Liu, Y., Patricelli, M.P. & Cravatt, B.F. Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* **96**, 14694-14699 (1999).
63. Lomenick, B., Jung, G., Wohlschlegel, J.A. & Huang, J. Target identification using drug affinity responsive target stability (DARTS). *Curr Protoc Chem Biol* **3**, 163-180 (2011).
64. Pai, M.Y. et al. Drug affinity responsive target stability (DARTS) for small-molecule target identification. *Methods Mol Biol* **1263**, 287-298 (2015).
65. Dal Piaz, F. et al. Drug affinity responsive target stability (DARTS) identifies laurifolioside as a new clathrin heavy chain modulator. *J Nat Prod* **79**, 2681-2692 (2016).
66. Chan, J.N. et al. Target identification by chromatographic co-elution: monitoring of drug-protein interactions without immobilization or chemical derivatization. *Mol Cell Proteomics* **11**, M111 016642 (2012).
67. Savitski, M.M. et al. Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, 1255784 (2014).
68. Martinez Molina, D. et al. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84-87 (2013).
69. Jafari, R. et al. The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat Protoc* **9**, 2100-2122 (2014).
70. Alshareef, A. et al. The use of cellular thermal shift assay (CETSA) to study crizotinib resistance in ALK-expressing human cancers. *Sci Rep* **6**, 33710 (2016).
71. Shaw, J. et al. Determining direct binders of the Androgen Receptor using a high-throughput Cellular Thermal Shift Assay. *Sci Rep* **8**, 163 (2018).
72. Franken, H. et al. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* **10**, 1567-1593 (2015).
73. Fomina-Yadlin, D. et al. Small-molecule inducers of insulin expression in pancreatic alpha-cells. *Proc Natl Acad Sci U S A* **107**, 15099-15104 (2010).
74. Fabian, M.A. et al. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat Biotechnol* **23**, 329-336 (2005).
75. Zhou, F.X., Bonin, J. & Predki, P.F. Development of functional protein microarrays for drug discovery: progress and challenges. *Comb Chem High Throughput Screen* **7**, 539-546 (2004).
76. Tu, S., Jiang, H.W., Liu, C.X., Zhou, S.M. & Tao, S.C. Protein microarrays for studies of drug mechanisms and biomarker discovery in the era of systems biology. *Curr Pharm Des* **20**, 49-55 (2014).
77. West, G.M., Tang, L. & Fitzgerald, M.C. Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass spectrometry-based strategy. *Anal Chem* **80**, 4175-4185 (2008).
78. Kurganov, B.I. Kinetics of protein aggregation. Quantitative estimation of the chaperone-like activity in test-systems based on suppression of protein aggregation. *Biochemistry (Mosc)* **67**, 409-422 (2002).

79. Vedadi, M. et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* **103**, 15835-15840 (2006).
80. Patching, S.G. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim Biophys Acta* **1838**, 43-55 (2014).
81. Zhu, J. et al. Protein interaction discovery using parallel analysis of translated ORFs (PLATO). *Nat Biotechnol* **31**, 331-334 (2013).
82. Larman, H.B., Liang, A.C., Elledge, S.J. & Zhu, J. Discovery of protein interactions using parallel analysis of translated ORFs (PLATO). *Nat Protoc* **9**, 90-103 (2014).
83. Salcius, M. et al. SEC-TID: a label-free method for small-molecule target identification. *J Biomol Screen* **19**, 917-927 (2014).
84. Degorce, F. et al. HTRF: A technology tailored for drug discovery - a review of theoretical aspects and recent applications. *Curr Chem Genomics* **3**, 22-32 (2009).
85. Bertucci, C., Pistolozzi, M. & De Simone, A. Circular dichroism in drug discovery and development: an abridged review. *Anal Bioanal Chem* **398**, 155-166 (2010).
86. Masson, G.R., Jenkins, M.L. & Burke, J.E. An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin Drug Discov* **12**, 981-994 (2017).
87. Wang, Y. et al. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* **37**, W623-633 (2009).
88. Johannessen, C.M., Clemons, P.A. & Wagner, B.K. Integrating phenotypic small-molecule profiling and human genetics: the next phase in drug discovery. *Trends Genet* **31**, 16-23 (2015).
89. Feng, Y., Mitchison, T.J., Bender, A., Young, D.W. & Tallarico, J.A. Multi-parameter phenotypic profiling: using cellular effects to characterize small-molecule compounds. *Nat Rev Drug Discov* **8**, 567-578 (2009).
90. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607 (2012).
91. Paull, K.D. et al. Display and analysis of patterns of differential activity of drugs against human tumor cell lines: development of mean graph and COMPARE algorithm. *J Natl Cancer Inst* **81**, 1088-1092 (1989).
92. Weinstein, J.N. et al. An information-intensive approach to the molecular pharmacology of cancer. *Science* **275**, 343-349 (1997).
93. Hughes, T.R. et al. Functional discovery via a compendium of expression profiles. *Cell* **102**, 109-126 (2000).
94. Marton, M.J. et al. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* **4**, 1293-1301 (1998).
95. Ganter, B. et al. Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *J Biotechnol* **119**, 219-244 (2005).
96. Lamb, J. et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929-1935 (2006).
97. Hieronymus, H. et al. Gene expression signature-based chemical genomic prediction identifies a novel class of HSP90 pathway modulators. *Cancer Cell* **10**, 321-330 (2006).
98. Cheng, T., Li, Q., Wang, Y. & Bryant, S.H. Identifying compound-target associations by combining bioactivity profile similarity search and public databases mining. *J Chem Inf Model* **51**, 2440-2448 (2011).
99. Perlman, Z.E. et al. Multidimensional drug profiling by automated microscopy. *Science* **306**, 1194-1198 (2004).
100. Carpenter, A.E. Image-based chemical screening. *Nat Chem Biol* **3**, 461-465 (2007).
101. Xia, X. et al. Image-based chemical screening identifies drug efflux inhibitors in lung cancer cells. *Cancer Res* **70**, 7723-7733 (2010).
102. Wolpaw, A.J. et al. Modulatory profiling identifies mechanisms of small molecule-induced cell death. *Proc Natl Acad Sci U S A* **108**, E771-780 (2011).
103. Laggner, C. et al. Chemical informatics and target identification in a zebrafish phenotypic screen. *Nat Chem Biol* **8**, 144-146 (2011).

104. Gregori-Puigjane, E. et al. Identifying mechanism-of-action targets for drugs and probes. *Proc Natl Acad Sci U S A* **109**, 11178-11183 (2012).
105. Lounkine, E. et al. Large-scale prediction and testing of drug activity on side-effect targets. *Nature* **486**, 361-367 (2012).
106. Keiser, M.J. et al. Predicting new molecular targets for known drugs. *Nature* **462**, 175-181 (2009).
107. Keiser, M.J. et al. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* **25**, 197-206 (2007).
108. Lagunin, A., Stepanchikova, A., Filimonov, D. & Poroikov, V. PASS: prediction of activity spectra for biologically active substances. *Bioinformatics* **16**, 747-748 (2000).
109. Anzali, S. et al. Discriminating between drugs and nondrugs by prediction of activity spectra for substances (PASS). *J Med Chem* **44**, 2432-2437 (2001).
110. Nidhi, Glick, M., Davies, J.W. & Jenkins, J.L. Prediction of biological targets for compounds using multiple-category Bayesian models trained on chemogenomics databases. *J Chem Inf Model* **46**, 1124-1133 (2006).
111. Martinez-Jimenez, F. et al. Target prediction for an open access set of compounds active against Mycobacterium tuberculosis. *PLoS Comput Biol* **9**, e1003253 (2013).
112. Chen, B., Ding, Y. & Wild, D.J. Assessing drug target association using semantic linked data. *PLoS Comput Biol* **8**, e1002574 (2012).
113. Andersson, C.R., Gustafsson, M.G. & Strombergsson, H. Quantitative chemogenomics: machine-learning models of protein-ligand interaction. *Curr Top Med Chem* **11**, 1978-1993 (2011).
114. Bender, A. et al. Use of ligand based models for protein domains to predict novel molecular targets and applications to triage affinity chromatography data. *J Proteome Res* **8**, 2575-2585 (2009).
115. Kufareva, I., Chen, Y.C., Ilatovskiy, A.V. & Abagyan, R. Compound activity prediction using models of binding pockets or ligand properties in 3D. *Curr Top Med Chem* **12**, 1869-1882 (2012).
116. Gao, Z. et al. PDTD: a web-accessible protein database for drug target identification. *BMC Bioinformatics* **9**, 104 (2008).
117. Chen, X., Ung, C.Y. & Chen, Y. Can an in silico drug-target search method be used to probe potential mechanisms of medicinal plant ingredients? *Nat Prod Rep* **20**, 432-444 (2003).
118. Kellenberger, E., Rodrigo, J., Muller, P. & Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **57**, 225-242 (2004).
119. Zheng, X.S., Chan, T.F. & Zhou, H.H. Genetic and genomic approaches to identify and study the targets of bioactive small molecules. *Chem Biol* **11**, 609-618 (2004).
120. Kasap, C., Elemento, O. & Kapoor, T.M. DrugTargetSeqR: a genomics- and CRISPR-Cas9-based method to analyze drug targets. *Nat Chem Biol* **10**, 626-628 (2014).
121. Smurnyy, Y. et al. DNA sequencing and CRISPR-Cas9 gene editing for target validation in mammalian cells. *Nat Chem Biol* **10**, 623-625 (2014).
122. Smith, A.M., Ammar, R., Nislow, C. & Giaever, G. A survey of yeast genomic assays for drug and target discovery. *Pharmacol Ther* **127**, 156-164 (2010).
123. Rine, J., Hansen, W., Hardeman, E. & Davis, R.W. Targeted selection of recombinant clones through gene dosage effects. *Proc Natl Acad Sci U S A* **80**, 6750-6754 (1983).
124. Hartwell, L.H., Szankasi, P., Roberts, C.J., Murray, A.W. & Friend, S.H. Integrating genetic approaches into the discovery of anticancer drugs. *Science* **278**, 1064-1068 (1997).
125. Perlstein, E.O., Ruderfer, D.M., Roberts, D.C., Schreiber, S.L. & Kruglyak, L. Genetic basis of individual differences in the response to small-molecule drugs in yeast. *Nat Genet* **39**, 496-502 (2007).
126. Launhardt, H., Hinnen, A. & Munder, T. Drug-induced phenotypes provide a tool for the functional analysis of yeast genes. *Yeast* **14**, 935-942 (1998).
127. Giaever, G. et al. Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat Genet* **21**, 278-283 (1999).
128. Giaever, G. et al. Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci U S A* **101**, 793-798 (2004).
129. Baetz, K. et al. Yeast genome-wide drug-induced haploinsufficiency screen to determine drug mode of action. *Proc Natl Acad Sci U S A* **101**, 4525-4530 (2004).

130. Shoemaker, D.D., Lashkari, D.A., Morris, D., Mittmann, M. & Davis, R.W. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* **14**, 450-456 (1996).
131. Pierce, S.E. et al. A unique and universal molecular barcode array. *Nat Methods* **3**, 601-603 (2006).
132. Lum, P.Y. et al. Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell* **116**, 121-137 (2004).
133. Tong, A.H. et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364-2368 (2001).
134. Baryshnikova, A. et al. Synthetic genetic array (SGA) analysis in Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Methods Enzymol* **470**, 145-179 (2010).
135. Schuldiner, M. et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**, 507-519 (2005).
136. Breslow, D.K. et al. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* **5**, 711-718 (2008).
137. Yan, Z. et al. Yeast Barcoders: a chemogenomic application of a universal donor-strain collection carrying bar-code identifiers. *Nat Methods* **5**, 719-725 (2008).
138. Mnaimneh, S. et al. Exploration of essential gene functions via titratable promoter alleles. *Cell* **118**, 31-44 (2004).
139. Kofoed, M. et al. An updated collection of sequence barcoded temperature-sensitive alleles of yeast essential genes. *G3 (Bethesda)* **5**, 1879-1887 (2015).
140. Li, Z. et al. Systematic exploration of essential yeast gene function with temperature-sensitive mutants. *Nat Biotechnol* **29**, 361-367 (2011).
141. Carroll, S.Y. et al. A yeast killer toxin screen provides insights into a/b toxin entry, trafficking, and killing mechanisms. *Dev Cell* **17**, 552-560 (2009).
142. Parsons, A.B. et al. Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611-625 (2006).
143. Lee, W. et al. Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet* **1**, e24 (2005).
144. Parsons, A.B. et al. Integration of chemical-genetic and genetic interaction data links bioactive compounds to cellular target pathways. *Nat Biotechnol* **22**, 62-69 (2004).
145. Butcher, R.A. et al. Microarray-based method for monitoring yeast overexpression strains reveals small-molecule targets in TOR pathway. *Nat Chem Biol* **2**, 103-109 (2006).
146. Luesch, H. et al. A genome-wide overexpression screen in yeast for small-molecule target identification. *Chem Biol* **12**, 55-63 (2005).
147. Ho, C.H. et al. A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. *Nat Biotechnol* **27**, 369-377 (2009).
148. Hoon, S. et al. An integrated platform of genomic assays reveals small-molecule bioactivities. *Nat Chem Biol* **4**, 498-506 (2008).
149. Heitman, J., Movva, N.R. & Hall, M.N. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science* **253**, 905-909 (1991).
150. Peyroche, A. et al. Brefeldin A acts to stabilize an abortive ARF-GDP-Sec7 domain protein complex: involvement of specific residues of the Sec7 domain. *Mol Cell* **3**, 275-285 (1999).
151. Ojini, I. & Gammie, A. Rapid identification of chemoresistance mechanisms using yeast DNA mismatch repair mutants. *G3 (Bethesda)* **5**, 1925-1935 (2015).
152. Huang, Z. et al. A functional variomics tool for discovering drug-resistance genes and drug targets. *Cell Rep* **3**, 577-585 (2013).
153. Venter, J.C. et al. The sequence of the human genome. *Science* **291**, 1304-1351 (2001).
154. Lander, E.S. et al. Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001).
155. Margulies, M. et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
156. Hu, X. & Zhang, Z. Understanding the genetic mechanisms of cancer drug resistance using genomic approaches. *Trends Genet* **32**, 127-137 (2016).
157. Bauer, J.A. et al. RNA interference (RNAi) screening approach identifies agents that enhance paclitaxel activity in breast cancer cells. *Breast Cancer Res* **12**, R41 (2010).

158. Mohr, S., Bakal, C. & Perrimon, N. Genomic screening with RNAi: results and challenges. *Annu Rev Biochem* **79**, 37-64 (2010).
159. Wilson, R.C. & Doudna, J.A. Molecular mechanisms of RNA interference. *Annu Rev Biophys* **42**, 217-239 (2013).
160. Westbrook, T.F. et al. A genetic screen for candidate tumor suppressors identifies REST. *Cell* **121**, 837-848 (2005).
161. Kolfschoten, I.G. et al. A genetic screen identifies PITX1 as a suppressor of RAS activity and tumorigenicity. *Cell* **121**, 849-858 (2005).
162. Putzbach, W. et al. Many si/shRNAs can kill cancer cells by targeting multiple survival genes through an off-target mechanism. *eLife* **6**, e29702 (2017).
163. Bassik, M.C. et al. A systematic mammalian genetic interaction map reveals pathways underlying ricin susceptibility. *Cell* **152**, 909-922 (2013).
164. Mohr, S.E., Smith, J.A., Shamu, C.E., Neumuller, R.A. & Perrimon, N. RNAi screening comes of age: improved techniques and complementary approaches. *Nat Rev Mol Cell Biol* **15**, 591-600 (2014).
165. Kaelin, W.G., Jr. Molecular biology. Use and abuse of RNAi to study mammalian gene function. *Science* **337**, 421-422 (2012).
166. Dai, Z. et al. edgeR: a versatile tool for the analysis of shRNA-seq and CRISPR-Cas9 genetic screens. *F1000Res* **3**, 95 (2014).
167. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
168. Matheny, C.J. et al. Next-generation NAMPT inhibitors identified by sequential high-throughput phenotypic chemical and functional genomic screens. *Chem Biol* **20**, 1352-1363 (2013).
169. Sidrauski, C. et al. Pharmacological brake-release of mRNA translation enhances cognitive memory. *eLife* **2**, e00498 (2013).
170. Brummelkamp, T.R. et al. An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. *Nat Chem Biol* **2**, 202-206 (2006).
171. Luo, B. et al. Highly parallel identification of essential genes in cancer cells. *Proc Natl Acad Sci U S A* **105**, 20380-20385 (2008).
172. Burgess, D.J. et al. Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A* **105**, 9053-9058 (2008).
173. Moffat, J. et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124**, 1283-1298 (2006).
174. Wang, J. et al. Cellular phenotype recognition for high-content RNA interference genome-wide screening. *J Biomol Screen* **13**, 29-39 (2008).
175. Eggert, U.S. et al. Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. *PLoS Biol* **2**, e379 (2004).
176. Smith, I. et al. Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. *PLoS Biol* **15**, e2003213 (2017).
177. Perreira, J.M., Meraner, P. & Brass, A.L. Functional genomic strategies for elucidating human-virus interactions: will CRISPR knockout RNAi and haploid cells? *Adv Virus Res* **94**, 1-51 (2016).
178. Unniyampurath, U., Pilankatta, R. & Krishnan, M.N. RNA interference in the age of CRISPR: will CRISPR interfere with RNAi? *Int J Mol Sci* **17**, 291 (2016).
179. Lee, T., Shah, C. & Xu, E.Y. Gene trap mutagenesis: a functional genomics approach towards reproductive research. *Mol Hum Reprod* **13**, 771-779 (2007).
180. Carette, J.E. et al. Global gene disruption in human cells to assign genes to phenotypes by deep sequencing. *Nat Biotechnol* **29**, 542-546 (2011).
181. Carette, J.E. et al. Haploid genetic screens in human cells identify host factors used by pathogens. *Science* **326**, 1231-1235 (2009).
182. DeNicola, G.M., Karreth, F.A., Adams, D.J. & Wong, C.C. The utility of transposon mutagenesis for cancer studies in the era of genome editing. *Genome Biol* **16**, 229 (2015).
183. Dupuy, A.J., Akagi, K., Largaespada, D.A., Copeland, N.G. & Jenkins, N.A. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. *Nature* **436**, 221-226 (2005).
184. Blomen, V.A. et al. Gene essentiality and synthetic lethality in haploid human cells. *Science* **350**, 1092-1096 (2015).
185. Wang, T. et al. Identification and characterization of essential genes in the human genome. *Science* **350**, 1096-1101 (2015).

186. Shen, H. et al. Identification of genes that nodulate susceptibility to formaldehyde and imatinib by functional genomic screening in human haploid KBM7 cells. *Toxicol Sci* **151**, 10-22 (2016).
187. Pettitt, S.J. et al. Genome-wide barcoded transposon screen for cancer drug sensitivity in haploid mouse embryonic stem cells. *Sci Data* **4**, 170020 (2017).
188. Dubey, R. et al. Chromatin-remodeling complex SWI/SNF controls multidrug resistance by transcriptionally regulating the drug efflux pump ABCB1. *Cancer Res* **76**, 5810-5821 (2016).
189. Winter, G.E. et al. The solute carrier SLC35F2 enables YM155-mediated DNA damage toxicity. *Nat Chem Biol* **10**, 768-773 (2014).
190. Heijink, A.M. et al. A haploid genetic screen identifies the G1/S regulatory machinery as a determinant of Wee1 inhibitor sensitivity. *Proc Natl Acad Sci U S A* **112**, 15160-15165 (2015).
191. Pettitt, S.J. et al. A genetic screen using the PiggyBac transposon in haploid cells identifies Parp1 as a mediator of olaparib toxicity. *PLoS One* **8**, e61520 (2013).
192. Reiling, J.H. et al. A haploid genetic screen identifies the major facilitator domain containing 2A (MFSD2A) transporter as a key mediator in the response to tunicamycin. *Proc Natl Acad Sci U S A* **108**, 11756-11765 (2011).
193. Arnoldo, A. et al. A genome scale overexpression screen to reveal drug activity in human cells. *Genome Med* **6**, 32 (2014).
194. Yang, X. et al. A public genome-scale lentiviral expression library of human ORFs. *Nat Methods* **8**, 659-661 (2011).
195. Wilson, F.H. et al. A functional landscape of resistance to ALK inhibition in lung cancer. *Cancer Cell* **27**, 397-408 (2015).
196. Iorio, F. et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc Natl Acad Sci U S A* **107**, 14621-14626 (2010).
197. Chen, L. et al. Transposon activation mutagenesis as a screening tool for identifying resistance to cancer therapeutics. *BMC Cancer* **13**, 93 (2013).
198. Rees, M.G. et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* **12**, 109-116 (2016).
199. Hamza, A. et al. Complementation of yeast genes with human genes as an experimental platform for functional testing of human genetic variants. *Genetics* **201**, 1263-1274 (2015).
200. Kapoor, T.M. & Miller, R.M. Leveraging chemotype-specific resistance for drug target identification and chemical biology. *Trends Pharmacol Sci* **38**, 1100-1109 (2017).
201. Gottesman, M.M., Fojo, T. & Bates, S.E. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* **2**, 48-58 (2002).
202. McDermott, M. et al. In vitro development of chemotherapy and targeted therapy drug-resistant cancer cell lines: a practical guide with case studies. *Front Oncol* **4**, 40 (2014).
203. Wacker, S.A., Houghtaling, B.R., Elemento, O. & Kapoor, T.M. Using transcriptome sequencing to identify mechanisms of drug action and resistance. *Nat Chem Biol* **8**, 235-237 (2012).
204. Brammeld, J.S. et al. Genome-wide chemical mutagenesis screens allow unbiased saturation of the cancer genome and identification of drug resistance mutations. *Genome Res* **27**, 613-625 (2017).
205. Forment, J.V. et al. Genome-wide genetic screening with chemically mutagenized haploid embryonic stem cells. *Nat Chem Biol* **13**, 12-14 (2017).
206. Khan, A.H., Bloom, J.S., Faridmoayer, E. & Smith, D.J. Genetic screening reveals a link between Wnt signaling and antitubulin drugs. *Pharmacogenomics J* **16**, 164-172 (2016).
207. Silva, G. et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther* **11**, 11-27 (2011).
208. Sander, J.D. & Joung, J.K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347-355 (2014).
209. Mansour, S.L., Thomas, K.R. & Capecchi, M.R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348-352 (1988).
210. Choulika, A., Perrin, A., Dujon, B. & Nicolas, J.F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol Cell Biol* **15**, 1968-1973 (1995).
211. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* **14**, 8096-8106 (1994).

212. Rudin, N., Sugarman, E. & Haber, J.E. Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae. *Genetics* **122**, 519-534 (1989).
213. Schmid-Burgk, J.L. Disruptive non-disruptive applications of CRISPR/Cas9. *Curr Opin Biotechnol* **48**, 203-209 (2017).
214. Wang, T., Wei, J.J., Sabatini, D.M. & Lander, E.S. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343**, 80-84 (2014).
215. Barrangou, R. & Horvath, P. A decade of discovery: CRISPR functions and applications. *Nat Microbiol* **2**, 17092 (2017).
216. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826 (2013).
217. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843 (2013).
218. Zischewski, J., Fischer, R. & Bortesi, L. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol Adv* **35**, 95-104 (2017).
219. Kleinstiver, B.P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495 (2016).
220. Maggio, I. & Goncalves, M.A. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol* **33**, 280-291 (2015).
221. Chen, J.S. et al. Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* **550**, 407-410 (2017).
222. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).
223. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232 (2013).
224. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).
225. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821 (2012).
226. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771 (2015).
227. Wright, A.V., Nunez, J.K. & Doudna, J.A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29-44 (2016).
228. van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M. & Brouns, S.J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* **34**, 401-407 (2009).
229. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712 (2007).
230. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S.D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551-2561 (2005).
231. Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174-182 (2005).
232. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653-663 (2005).
233. Koonin, E.V., Makarova, K.S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* **37**, 67-78 (2017).
234. Makarova, K.S. et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722-736 (2015).
235. Mohanraju, P. et al. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
236. Charpentier, E., Richter, H., van der Oost, J. & White, M.F. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiol Rev* **39**, 428-441 (2015).
237. Burstein, D. et al. New CRISPR-Cas systems from uncultivated microbes. *Nature* **542**, 237-241 (2017).
238. Smargon, A.A. et al. Cas13b Is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol Cell* **65**, 618-630 e617 (2017).

239. Abudayyeh, O.O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
240. Nishimasu, H. et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935-949 (2014).
241. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569-573 (2014).
242. Jinek, M. et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014).
243. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. & Joung, J.K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279-284 (2014).
244. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827-832 (2013).
245. Jasin, M. & Haber, J.E. The democratization of gene editing: Insights from site-specific cleavage and double-strand break repair. *DNA Repair (Amst)* **44**, 6-16 (2016).
246. Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773-782 (2011).
247. Chapman, J.R., Taylor, M.R. & Boulton, S.J. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell* **47**, 497-510 (2012).
248. Ran, F.A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).
249. Dominguez, A.A., Lim, W.A. & Qi, L.S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* **17**, 5-15 (2016).
250. Hess, G.T., Tycko, J., Yao, D. & Bassik, M.C. Methods and applications of CRISPR-mediated base editing in eukaryotic genomes. *Mol Cell* **68**, 26-43 (2017).
251. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
252. Chen, B. et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* **155**, 1479-1491 (2013).
253. Joung, J. et al. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat Protoc* **12**, 828-863 (2017).
254. Burckstummer, T. et al. A reversible gene trap collection empowers haploid genetics in human cells. *Nat Methods* **10**, 965-971 (2013).
255. Carette, J.E. et al. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* **477**, 340-343 (2011).
256. Chidley, C., Trauger, S.A., Birsoy, K. & O'Shea, E.K. The anticancer natural product ophiobolin A induces cytotoxicity by covalent modification of phosphatidylethanolamine. *eLife* **5,** e14601 (2016).
257. Koike-Yusa, H., Li, Y., Tan, E.P., Velasco-Herrera Mdel, C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol* **32**, 267-273 (2014).
258. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).
259. Aguirre, A.J. et al. Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. *Cancer Discov* **6**, 914-929 (2016).
260. Meyers, R.M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat Genet* **49**, 1779-1784 (2017).
261. Doench, J.G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191 (2016).
262. Hart, T. et al. Evaluation and design of genome-wide CRISPR/SpCas9 knockout Screens. *G3 (Bethesda)* **7**, 2719-2727 (2017).
263. Timms, R.T. et al. Genetic dissection of mammalian ERAD through comparative haploid and CRISPR forward genetic screens. *Nat Commun* **7**, 11786 (2016).
264. Moder, M. et al. Parallel genome-wide screens identify synthetic viable interactions between the BLM helicase complex and Fanconi anemia. *Nat Commun* **8**, 1238 (2017).
265. Deans, R.M. et al. Parallel shRNA and CRISPR-Cas9 screens enable antiviral drug target identification. *Nat Chem Biol* **12**, 361-366 (2016).

266. Evers, B. et al. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotechnol* **34**, 631-633 (2016).
267. Morgens, D.W., Deans, R.M., Li, A. & Bassik, M.C. Systematic comparison of CRISPR/Cas9 and RNAi screens for essential genes. *Nat Biotechnol* **34**, 634-636 (2016).
268. Bertomeu, T. et al. A high-resolution genome-wide CRISPR/Cas9 viability screen reveals structural features and contextual diversity of the human cell-essential proteome. *Mol Cell Biol* **38** (2018).
269. Hart, T. et al. High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell* **163**, 1515-1526 (2015).
270. Wang, T. et al. Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. *Cell* **168**, 890-903 e815 (2017).
271. Kodama, M. et al. In vivo loss-of-function screens identify KPNB1 as a new druggable oncogene in epithelial ovarian cancer. *Proc Natl Acad Sci U S A* **114**, E7301-E7310 (2017).
272. Fei, T. et al. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc Natl Acad Sci U S A* **114**, E5207-E5215 (2017).
273. Tzelepis, K. et al. A CRISPR dropout screen Identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep* **17**, 1193-1205 (2016).
274. Kiessling, M.K. et al. Identification of oncogenic driver mutations by genome-wide CRISPR-Cas9 dropout screening. *BMC Genomics* **17**, 723 (2016).
275. Rosenbluh, J. et al. Complementary information derived from CRISPR Cas9 mediated gene deletion and suppression. *Nat Commun* **8**, 15403 (2017).
276. Parnas, O. et al. A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. *Cell* **162**, 675-686 (2015).
277. Zhou, Y. et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature* **509**, 487-491 (2014).
278. Virreira Winter, S., Zychlinsky, A. & Bardoel, B.W. Genome-wide CRISPR screen reveals novel host factors required for Staphylococcus aureus alpha-hemolysin-mediated toxicity. *Sci Rep* **6**, 24242 (2016).
279. Shi, C.X. et al. CRISPR genome-wide screening identifies dependence on the proteasome subunit PSMC6 for bortezomib sensitivity in multiple myeloma. *Mol Cancer Ther* **16**, 2862-2870 (2017).
280. Krall, E.B. et al. KEAP1 loss modulates sensitivity to kinase targeted therapy in lung cancer. e*Life* **6,** e18970 (2017).
281. Hou, P. et al. A genome-wide CRISPR screen identifies genes critical for resistance to FLT3 inhibitor AC220. *Cancer Res* **77**, 4402-4413 (2017).
282. Kurata, M. et al. Using genome-wide CRISPR library screening with library resistant DCK to find new sources of Ara-C drug resistance in AML. *Sci Rep* **6**, 36199 (2016).
283. Tao, L. et al. Frizzled proteins are colonic epithelial receptors for C. difficile toxin B. *Nature* **538**, 350-355 (2016).
284. Larson, M.H. et al. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **8**, 2180-2196 (2013).
285. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
286. Horlbeck, M.A. et al. Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5,** e19760 (2016).
287. Horlbeck, M.A. et al. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *eLife* **5,** e12677 (2016).
288. Gilbert, L.A. et al. Genome-scale CRISPR-mediated control of gene repression and activation. *Cell* **159**, 647-661 (2014).
289. Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).
290. Jost, M. et al. Combined CRISPRi/a-based chemical genetic screens reveal that rigosertib is a microtubule-destabilizing agent. *Mol Cell* **68**, 210-223 e216 (2017).
291. Anderson, D.J. et al. Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. *Cancer Cell* **28**, 653-665 (2015).
292. Smith, J.D. et al. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol* **17**, 45 (2016).

293. Heaton, B.E. et al. A CRISPR activation screen identifies a pan-avian influenza virus inhibitory host factor. *Cell Rep* **20**, 1503-1512 (2017).
294. le Sage, C. et al. Dual direction CRISPR transcriptional regulation screening uncovers gene networks driving drug resistance. *Sci Rep* **7**, 17693 (2017).
295. Du, D. et al. Genetic interaction mapping in mammalian cells using CRISPR interference. *Nat Methods* **14**, 577-580 (2017).
296. Liu, X. et al. High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. *Mol Syst Biol* **13**, 931 (2017).
297. Peters, J.M. et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493-1506 (2016).
298. Joung, J. et al. Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. *Nature* **548**, 343-346 (2017).
299. Klann, T.S. et al. CRISPR-Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nat Biotechnol* **35**, 561-568 (2017).
300. Liu, S.J. et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355** (2017).
301. Adamson, B. et al. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* **167**, 1867-1882 e1821 (2016).
302. Auerbach, C., Robson, J.M. & Carr, J.G. The chemical production of mutations. *Science* **105**, 243-247 (1947).
303. Hess, G.T. et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat Methods* **13**, 1036-1042 (2016).
304. Ma, Y. et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods* **13**, 1029-1035 (2016).
305. Nishida, K. et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353,** aaf8729 (2016).
306. Fellmann, C., Gowen, B.G., Lin, P.C., Doudna, J.A. & Corn, J.E. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat Rev Drug Discov* **16**, 89-100 (2017).
307. Dai, W.J. et al. CRISPR-Cas9 for in vivo gene therapy: promise and hurdles. *Mol Ther Nucleic Acids* **5**, e349 (2016).
308. Cyranoski, D. Chinese scientists to pioneer first human CRISPR trial. *Nature* **535**, 476-477 (2016).
309. Baltimore, D. et al. Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* **348**, 36-38 (2015).
310. Kohn, D.B., Porteus, M.H. & Scharenberg, A.M. Ethical and regulatory aspects of genome editing. *Blood* **127**, 2553-2560 (2016).
311. Esvelt, K.M., Smidler, A.L., Catteruccia, F. & Church, G.M. Concerning RNA-guided gene drives for the alteration of wild populations. e*Life* **3,** e03401 (2014).
312. Hammond, A. et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nat Biotechnol* **34**, 78-83 (2016).
313. Gantz, V.M. et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. *Proc Natl Acad Sci U S A* **112**, E6736-6743 (2015).
314. Proudfoot, C. et al. Genome edited sheep and cattle. *Transgenic Res* **24**, 147-153 (2015).
315. Carroll, D. Genome Editing: past, present, and future. *Yale J Biol Med* **90**, 653-659 (2017).
316. Plummer, R.J., Guo, Y. & Peng, Y. A CRISPR reimagining: New twists and turns of CRISPR beyond the genome-engineering revolution. *J Cell Biochem* **119**, 1299-1308 (2018).
317. Kwon, D.Y., Zhao, Y.T., Lamonica, J.M. & Zhou, Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat Commun* **8**, 15315 (2017).
318. Vora, S., Tuttle, M., Cheng, J. & Church, G. Next stop for the CRISPR revolution: RNA-guided epigenetic regulators. *FEBS J* **283**, 3181-3193 (2016).
319. Mishra, A. & Hawkins, R.D. Three-dimensional genome architecture and emerging technologies: looping in disease. *Genome Med* **9**, 87 (2017).
320. Jusiak, B., Cleto, S., Perez-Pinera, P. & Lu, T.K. Engineering synthetic gene circuits in living cells with CRISPR technology. *Trends Biotechnol* **34**, 535-547 (2016).
321. Nielsen, A.A. & Voigt, C.A. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol* **10**, 763 (2014).

322. Shipman, S.L., Nivala, J., Macklis, J.D. & Church, G.M. CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature* **547**, 345-349 (2017).
323. Lusk, C.P. & King, M.C. The nucleus: keeping it together by keeping it apart. *Curr Opin Cell Biol* **44**, 44-50 (2017).
324. Tran, E.J., King, M.C. & Corbett, A.H. Macromolecular transport between the nucleus and the cytoplasm: Advances in mechanism and emerging links to disease. *Biochim Biophys Acta* **1843**, 2784-2795 (2014).
325. Hayama, R., Rout, M.P. & Fernandez-Martinez, J. The nuclear pore complex core scaffold and permeability barrier: variations of a common theme. *Curr Opin Cell Biol* **46**, 110-118 (2017).
326. Beck, M. & Hurt, E. The nuclear pore complex: understanding its function through structural insight. *Nat Rev Mol Cell Biol* **18**, 73-89 (2017).
327. Cagatay, T. & Chook, Y.M. Karyopherins in cancer. *Curr Opin Cell Biol* **52**, 30-42 (2018).
328. Chook, Y.M. & Suel, K.E. Nuclear import by karyopherin-betas: recognition and inhibition. *Biochim Biophys Acta* **1813**, 1593-1606 (2011).
329. O'Reilly, A.J., Dacks, J.B. & Field, M.C. Evolution of the karyopherin-beta family of nucleocytoplasmic transport factors; ancient origins and continued specialization. *PLoS One* **6**, e19308 (2011).
330. Miyamoto, Y., Yamada, K. & Yoneda, Y. Importin alpha: a key molecule in nuclear transport and non-transport functions. *J Biochem* **160**, 69-75 (2016).
331. Xu, D., Farmer, A. & Chook, Y.M. Recognition of nuclear targeting signals by Karyopherin-beta proteins. *Curr Opin Struct Biol* **20**, 782-790 (2010).
332. Dong, X. et al. Structural basis for leucine-rich nuclear export signal recognition by CRM1. *Nature* **458**, 1136-1141 (2009).
333. Guttler, T. & Gorlich, D. Ran-dependent nuclear export mediators: a structural perspective. *EMBO J* **30**, 3457-3474 (2011).
334. Lim, R.Y., Huang, B. & Kapinos, L.E. How to operate a nuclear pore complex by Kap-centric control. *Nucleus* **6**, 366-372 (2015).
335. Thakar, K., Karaca, S., Port, S.A., Urlaub, H. & Kehlenbach, R.H. Identification of CRM1-dependent nuclear export cargos using quantitative mass spectrometry. *Mol Cell Proteomics* **12**, 664-678 (2013).
336. Xu, D., Grishin, N.V. & Chook, Y.M. NESdb: a database of NES-containing CRM1 cargoes. *Mol Biol Cell* **23**, 3673-3676 (2012).
337. Fu, S.C., Huang, H.C., Horton, P. & Juan, H.F. ValidNESs: a database of validated leucine-rich nuclear export signals. *Nucleic Acids Res* **41**, D338-343 (2013).
338. Fung, H.Y., Fu, S.C., Brautigam, C.A. & Chook, Y.M. Structural determinants of nuclear export signal orientation in binding to exportin CRM1. *eLife* **4** (2015).
339. Kirli, K. et al. A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *eLife* **4,** e11466 (2015).
340. Fukuda, M. et al. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* **390**, 308-311 (1997).
341. Neville, M., Stutz, F., Lee, L., Davis, L.I. & Rosbash, M. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr Biol* **7**, 767-775 (1997).
342. Ossareh-Nazari, B., Bachelerie, F. & Dargemont, C. Evidence for a role of CRM1 in signal-mediated nuclear protein export. *Science* **278**, 141-144 (1997).
343. Fornerod, M., Ohno, M., Yoshida, M. & Mattaj, I.W. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* **90**, 1051-1060 (1997).
344. Stade, K., Ford, C.S., Guthrie, C. & Weis, K. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* **90**, 1041-1050 (1997).
345. Molina-Navarro, M.M. et al. Functional networks of nucleocytoplasmic transport-related genes differentiate ischemic and dilated cardiomyopathies. A new therapeutic opportunity. *PLoS One* **9**, e104709 (2014).
346. Boeynaems, S., Bogaert, E., Van Damme, P. & Van Den Bosch, L. Inside out: the role of nucleocytoplasmic transport in ALS and FTLD. *Acta Neuropathol* **132**, 159-173 (2016).

347. Chan, W.M. et al. Expanded polyglutamine domain possesses nuclear export activity which modulates subcellular localization and toxicity of polyQ disease protein via exportin-1. *Hum Mol Genet* **20**, 1738-1750 (2011).
348. Audsley, M.D., Jans, D.A. & Moseley, G.W. Roles of nuclear trafficking in infection by cytoplasmic negative-strand RNA viruses: paramyxoviruses and beyond. *J Gen Virol* **97**, 2463-2481 (2016).
349. Lakdawala, S.S., Fodor, E. & Subbarao, K. Moving on out: transport and packaging of influenza viral RNA into virions. *Annu Rev Virol* **3**, 411-427 (2016).
350. Wulan, W.N., Heydet, D., Walker, E.J., Gahan, M.E. & Ghildyal, R. Nucleocytoplasmic transport of nucleocapsid proteins of enveloped RNA viruses. *Front Microbiol* **6**, 553 (2015).
351. Stake, M.S., Bann, D.V., Kaddis, R.J. & Parent, L.J. Nuclear trafficking of retroviral RNAs and Gag proteins during late steps of replication. *Viruses* **5**, 2767-2795 (2013).
352. Schmid, M., Gonzalez, R.A. & Dobner, T. CRM1-dependent transport supports cytoplasmic accumulation of adenoviral early transcripts. *J Virol* **86**, 2282-2292 (2012).
353. Nguyen, K.T., Holloway, M.P. & Altura, R.A. The CRM1 nuclear export protein in normal development and disease. *Int J Biochem Mol Biol* **3**, 137-151 (2012).
354. Fannemel, M. et al. Haploinsufficiency of XPO1 and USP34 by a de novo 230 kb deletion in 2p15, in a patient with mild intellectual disability and cranio-facial dysmorphisms. *Eur J Med Genet* **57**, 513-519 (2014).
355. Vercruysse, T. & Daelemans, D. HIV-1 Rev multimerization: mechanism and insights. *Curr HIV Res* **11**, 623-634 (2013).
356. Haines, J.D. et al. Nuclear export inhibitors avert progression in preclinical models of inflammatory demyelination. *Nat Neurosci* **18**, 511-520 (2015).
357. Shen, A. et al. Expression of CRM1 in human gliomas and its significance in p27 expression and clinical prognosis. *Neurosurgery* **65**, 153-159; discussion 159-160 (2009).
358. Jain, P. et al. Clinical and molecular characteristics of XPO1 mutations in patients with chronic lymphocytic leukemia. *Am J Hematol* **91**, E478-E479 (2016).
359. Etchin, J. et al. KPT-330 inhibitor of CRM1 (XPO1)-mediated nuclear export has selective anti-leukaemic activity in preclinical models of T-cell acute lymphoblastic leukaemia and acute myeloid leukaemia. *Br J Haematol* **161**, 117-127 (2013).
360. Yao, Y. et al. The expression of CRM1 is associated with prognosis in human osteosarcoma. *Oncol Rep* **21**, 229-235 (2009).
361. van der Watt, P.J. et al. The Karyopherin proteins, Crm1 and Karyopherin beta1, are overexpressed in cervical cancer and are critical for cancer cell survival and proliferation. *Int J Cancer* **124**, 1829-1840 (2009).
362. Noske, A. et al. Expression of the nuclear export protein chromosomal region maintenance/exportin 1/XPO1 is a prognostic factor in human ovarian cancer. *Cancer* **112**, 1733-1743 (2008).
363. Xie, Q.L., Liu, Y. & Zhu, Y. Chromosome region maintenance 1 expression and its association with clinical pathological features in primary carcinoma of the liver. *Exp Ther Med* **12**, 59-68 (2016).
364. Camus, V., Miloudi, H., Taly, A., Sola, B. & Jardin, F. XPO1 in B cell hematological malignancies: from recurrent somatic mutations to targeted therapy. *J Hematol Oncol* **10**, 47 (2017).
365. Jardin, F. et al. Recurrent mutations of the exportin 1 gene (XPO1) and their impact on selective inhibitor of nuclear export compounds sensitivity in primary mediastinal B-cell lymphoma. *Am J Hematol* **91**, 923-930 (2016).
366. Camus, V. et al. Detection and prognostic value of recurrent exportin 1 mutations in tumor and cell-free circulating DNA of patients with classical Hodgkin lymphoma. *Haematologica* **101**, 1094-1101 (2016).
367. Puente, X.S. et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**, 101-105 (2011).
368. Lin, D.C. et al. Genomic and molecular characterization of esophageal squamous cell carcinoma. *Nat Genet* **46**, 467-473 (2014).
369. Kim, J. et al. XPO1-dependent nuclear export is a druggable vulnerability in KRAS-mutant lung cancer. *Nature* **538**, 114-117 (2016).
370. Hong, A.L. et al. Integrated genetic and pharmacologic interrogation of rare cancers. *Nat Commun* **7**, 11987 (2016).

371. Yoshimura, S.H. & Hirano, T. HEAT repeats - versatile arrays of amphiphilic helices working in crowded environments? *J Cell Sci* **129**, 3963-3970 (2016).
372. Monecke, T. et al. Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* **324**, 1087-1091 (2009).
373. Guttler, T. et al. NES consensus redefined by structures of PKI-type and Rev-type nuclear export signals bound to CRM1. *Nat Struct Mol Biol* **17**, 1367-1376 (2010).
374. Monecke, T., Dickmanns, A. & Ficner, R. Allosteric control of the exportin CRM1 unraveled by crystal structure analysis. *FEBS J* **281**, 4179-4194 (2014).
375. Langer, K., Dian, C., Rybin, V., Muller, C.W. & Petosa, C. Insights into the function of the CRM1 cofactor RanBP3 from the structure of its Ran-binding domain. *PLoS One* **6**, e17011 (2011).
376. Boons, E. et al. Human exportin-1 is a target for combined therapy of HIV and AIDS related lymphoma. *EBioMedicine* **2**, 1102-1113 (2015).
377. Askjaer, P., Jensen, T.H., Nilsson, J., Englmeier, L. & Kjems, J. The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *J Biol Chem* **273**, 33414-33422 (1998).
378. Kimura, M. & Imamoto, N. Biological significance of the importin-beta family-dependent nucleocytoplasmic transport pathways. *Traffic* **15**, 727-748 (2014).
379. Poon, I.K. & Jans, D.A. Regulation of nuclear transport: central role in development and transformation? *Traffic* **6**, 173-186 (2005).
380. Adachi, Y. & Yanagida, M. Higher order chromosome structure is affected by cold-sensitive mutations in a Schizosaccharomyces pombe gene crm1+ which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *J Cell Biol* **108**, 1195-1207 (1989).
381. Fornerod, M., van Baal, S., Valentine, V., Shapiro, D.N. & Grosveld, G. Chromosomal localization of genes encoding CAN/Nup214-interacting proteins--human CRM1 localizes to 2p16, whereas Nup88 localizes to 17p13 and is physically linked to SF2p32. *Genomics* **42**, 538-540 (1997).
382. Fornerod, M. et al. The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88. *EMBO J* **16**, 807-816 (1997).
383. Hamamoto, T., Uozumi, T. & Beppu, T. Leptomycins A and B, new antifungal antibiotics. III. Mode of action of leptomycin B on Schizosaccharomyces pombe. *J Antibiot (Tokyo)* **38**, 1573-1580 (1985).
384. Nishi, K. et al. Leptomycin B targets a regulatory cascade of crm1, a fission yeast nuclear protein, involved in control of higher order chromosome structure and gene expression. *J Biol Chem* **269**, 6320-6324 (1994).
385. Hamamoto, T., Gunji, S., Tsuji, H. & Beppu, T. Leptomycins A and B, new antifungal antibiotics. I. Taxonomy of the producing strain and their fermentation, purification and characterization. *J Antibiot (Tokyo)* **36**, 639-645 (1983).
386. Sun, Q. et al. Nuclear export inhibition through covalent conjugation and hydrolysis of Leptomycin B by CRM1. *Proc Natl Acad Sci U S A* **110**, 1303-1308 (2013).
387. Kudo, N. et al. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci U S A* **96**, 9112-9117 (1999).
388. Kudo, N. et al. Leptomycin B inhibition of signal-mediated nuclear export by direct binding to CRM1. *Exp Cell Res* **242**, 540-547 (1998).
389. Sun, Q. et al. Inhibiting cancer cell hallmark features through nuclear export inhibition. *Signal Transduct Target Ther* **1**, 16010 (2016).
390. Senapedis, W.T., Baloglu, E. & Landesman, Y. Clinical translation of nuclear export inhibitors in cancer. *Semin Cancer Biol* **27**, 74-86 (2014).
391. Komiyama, K. et al. Antitumor activity of leptomycin B. *J Antibiot (Tokyo)* **38**, 427-429 (1985).
392. Tunac, J.B., Graham, B.D., Dobson, W.E. & Lenzini, M.D. Novel antitumor antibiotics, CI-940 (PD 114,720) and PD 114,721. Taxonomy, fermentation and biological activity. *J Antibiot (Tokyo)* **38**, 460-465 (1985).
393. Burzlaff, A., Kalesse, M., Kasper, C. & Scheper, T. Multi parameter in vitro testing of ratjadone using flow cytometry. *Appl Microbiol Biotechnol* **62**, 174-179 (2003).
394. Newlands, E.S., Rustin, G.J. & Brampton, M.H. Phase I trial of elactocin. *Br J Cancer* **74**, 648-649 (1996).
395. Gademann, K. Controlling protein transport by small molecules. *Curr Drug Targets* **12**, 1574-1580 (2011).

396. Daelemans, D. et al. A synthetic HIV-1 Rev inhibitor interfering with the CRM1-mediated nuclear export. *Proc Natl Acad Sci U S A* **99**, 14440-14445 (2002).
397. Van Neck, T. et al. Inhibition of the CRM1-mediated nucleocytoplasmic transport by N-azolylacrylates: structure-activity relationship and mechanism of action. *Bioorg Med Chem* **16**, 9487-9497 (2008).
398. Kalid, O., Toledo Warshaviak, D., Shechter, S., Sherman, W. & Shacham, S. Consensus Induced Fit Docking (cIFD): methodology, validation, and application to the discovery of novel Crm1 inhibitors. *J Comput Aided Mol Des* **26**, 1217-1228 (2012).
399. Lapalombella, R. et al. Selective inhibitors of nuclear export show that CRM1/XPO1 is a target in chronic lymphocytic leukemia. *Blood* **120**, 4621-4634 (2012).
400. Ranganathan, P. et al. Preclinical activity of a novel CRM1 inhibitor in acute myeloid leukemia. *Blood* **120**, 1765-1773 (2012).
401. Walker, C.J. et al. Preclinical and clinical efficacy of XPO1/CRM1 inhibition by the karyopherin inhibitor KPT-330 in Ph+ leukemias. *Blood* **122**, 3034-3044 (2013).
402. Inoue, H. et al. CRM1 blockade by selective inhibitors of nuclear export attenuates kidney cancer growth. *J Urol* **189**, 2317-2326 (2013).
403. Zhang, K. et al. Novel selective inhibitors of nuclear export CRM1 antagonists for therapy in mantle cell lymphoma. *Exp Hematol* **41**, 67-78 e64 (2013).
404. Etchin, J. et al. Antileukemic activity of nuclear export inhibitors that spare normal hematopoietic cells. *Leukemia* **27**, 66-74 (2013).
405. Kojima, K. et al. Prognostic impact and targeting of CRM1 in acute myeloid leukemia. *Blood* **121**, 4166-4174 (2013).
406. Cheng, Y. et al. XPO1 (CRM1) inhibition represses STAT3 activation to drive a survivin-dependent oncogenic switch in triple-negative breast cancer. *Mol Cancer Ther* **13**, 675-686 (2014).
407. Tai, Y.T. et al. CRM1 inhibition induces tumor cell cytotoxicity and impairs osteoclastogenesis in multiple myeloma: molecular mechanisms and therapeutic implications. *Leukemia* **28**, 155-165 (2014).
408. Chen, C. et al. in EHA Annual Meeting (2014).
409. Mahaseth, H. et al. in ESMO Annual Meeting (Madrid; 2014).
410. Turner, J.G., Dawson, J. & Sullivan, D.M. Nuclear export of proteins and drug resistance in cancer. *Biochem Pharmacol* **83**, 1021-1032 (2012).
411. Kau, T.R., Way, J.C. & Silver, P.A. Nuclear transport and cancer: from mechanism to intervention. *Nat Rev Cancer* **4**, 106-117 (2004).
412. Aloisi, A. et al. BCR-ABL nuclear entrapment kills human CML cells: ex vivo study on 35 patients with the combination of imatinib mesylate and leptomycin B. *Blood* **107**, 1591-1598 (2006).
413. Zhang, Y. & Xiong, Y. A p53 amino-terminal nuclear export signal inhibited by DNA damage-induced phosphorylation. *Science* **292**, 1910-1915 (2001).
414. Falini, B. et al. Both carboxy-terminus NES motif and mutated tryptophan(s) are crucial for aberrant nuclear export of nucleophosmin leukemic mutants in NPMc+ AML. *Blood* **107**, 4514-4523 (2006).
415. Crochiere, M. et al. Deciphering mechanisms of drug sensitivity and resistance to Selective Inhibitor of Nuclear Export (SINE) compounds. *BMC Cancer* **15**, 910 (2015).
416. Daelemans, D., Costes, S.V., Lockett, S. & Pavlakis, G.N. Kinetic and molecular analysis of nuclear export factor CRM1 association with its cargo in vivo. *Mol Cell Biol* **25**, 728-739 (2005).
417. Costes, S.V. et al. Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophys J* **86**, 3993-4003 (2004).
418. Luthra, R. et al. Next-generation sequencing-based multigene mutational screening for acute myeloid leukemia using MiSeq: applicability for diagnostics and disease monitoring. *Haematologica* **99**, 465-473 (2014).
419. Kuruvilla, J. et al. in 55th ASH Annual Meeting and Exposition (2013).
420. Neggers, J.E. et al. Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing. *Chem Biol* **22**, 107-116 (2015).
421. Jeromin, S. et al. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia* **28**, 108-117 (2014).

422. Bond, J. et al. Cryptic XPO1-MLLT10 translocation is associated with HOXA locus deregulation in T-ALL. *Blood* **124**, 3023-3025 (2014).
423. Ran, F.A., Hsu, P.D., Wright, J., Agarwaka, V., Scott, D.A. & Zhang, F. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
424. Hess, D.T., Matsumoto, A., Kim, S.O., Marshall, H.E. & Stamler, J.S. Protein S-nitrosylation: purview and parameters. *Nat Rev Mol Cell Biol* **6**, 150-166 (2005).
425. Gould, N., Doulias, P.T., Tenopoulou, M., Raju, K. & Ischiropoulos, H. Regulation of protein function and signaling by reversible cysteine S-nitrosylation. *J Biol Chem* **288**, 26473-26479 (2013).
426. Smotrys, J.E. & Linder, M.E. Palmitoylation of intracellular signaling proteins: regulation and function. *Annu Rev Biochem* **73**, 559-587 (2004).
427. Lo Conte, M. & Carroll, K.S. The redox biochemistry of protein sulfenylation and sulfinylation. *J Biol Chem* **288**, 26480-26488 (2013).
428. Wang, P. et al. Repression of classical nuclear export by S-nitrosylation of CRM1. *J Cell Sci* **122**, 3772-3779 (2009).
429. Birsoy, K. et al. MCT1-mediated transport of a toxic molecule is an effective strategy for targeting glycolytic tumors. *Nat Genet* **45**, 104-108 (2013).
430. Greinert, R. et al. UVA-induced DNA double-strand breaks result from the repair of clustered oxidative DNA damages. *Nucleic Acids Res* **40**, 10263-10273 (2012).
431. Rath, O. & Kozielski, F. Kinesins and cancer. *Nat Rev Cancer* **12**, 527-539 (2012).
432. Titov, D.V. et al. XPB, a subunit of TFIIH, is a target of the natural product triptolide. *Nat Chem Biol* **7**, 182-188 (2011).
433. Essletzbichler, P. et al. Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. *Genome Res* **24**, 2059-2065 (2014).
434. Ipsaro, J.J. et al. Rapid generation of drug-resistance alleles at endogenous loci using CRISPR-Cas9 indel mutagenesis. *PLoS One* **12**, e0172177 (2017).
435. Donovan, K.F. et al. Creation of novel protein variants with CRISPR/Cas9-mediated mutagenesis: turning a screening by-product into a discovery tool. *PLoS One* **12**, e0170445 (2017).
436. Lu, S. & Wang, J. The resistance mechanisms of proteasome inhibitor bortezomib. *Biomark Res* **1**, 13 (2013).
437. Schrader, J. et al. The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* **353**, 594-598 (2016).
438. Fulciniti, M. et al. Functional role and therapeutic targeting of p21-activated kinase 4 in multiple myeloma. *Blood* **129**, 2233-2245 (2017).
439. Aboukameel, A. et al. Novel p21-activated kinase 4 (PAK4) allosteric modulators overcome drug resistance and stemness in pancreatic ductal adenocarcinoma. *Mol Cancer Ther* **16**, 76-87 (2017).
440. Khan, J.A., Tao, X. & Tong, L. Molecular basis for the inhibition of human NMPRTase, a novel target for anticancer agents. *Nat Struct Mol Biol* **13**, 582-588 (2006).
441. Watson, M. et al. The small molecule GMX1778 is a potent inhibitor of NAD+ biosynthesis: strategy for enhanced therapy in nicotinic acid phosphoribosyltransferase 1-deficient tumors. *Mol Cell Biol* **29**, 5872-5888 (2009).
442. Bi, T.Q. & Che, X.M. Nampt/PBEF/visfatin and cancer. *Cancer Biol Ther* **10**, 119-125 (2010).
443. Shackelford, R.E., Mayhall, K., Maxwell, N.M., Kandil, E. & Coppola, D. Nicotinamide phosphoribosyltransferase in malignancy: a review. *Genes Cancer* **4**, 447-456 (2013).
444. Hasmann, M. & Schemainda, I. FK866, a highly specific noncompetitive inhibitor of nicotinamide phosphoribosyltransferase, represents a novel mechanism for induction of tumor cell apoptosis. *Cancer Res* **63**, 7436-7442 (2003).
445. Sampath, D., Zabka, T.S., Misner, D.L., O'Brien, T. & Dragovich, P.S. Inhibition of nicotinamide phosphoribosyltransferase (NAMPT) as a therapeutic strategy in cancer. *Pharmacol Ther* **151**, 16-31 (2015).
446. Hovstadius, P. et al. A Phase I study of CHS 828 in patients with solid tumor malignancy. *Clin Cancer Res* **8**, 2843-2850 (2002).
447. Abu Aboud, O. et al. Dual and specific inhibition of NAMPT and PAK4 by KPT-9274 decreases kidney cancer growth. *Mol Cancer Ther* **15**, 2119-2129 (2016).

448. Jiang, Y.Y. et al. Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. *Gut* **66**, 1358-1368 (2017).
449. Lamb, J. The Connectivity Map: a new tool for biomedical research. *Nat Rev Cancer* **7**, 54-60 (2007).
450. Schmid-Burgk, J.L., Honing, K., Ebert, T.S. & Hornung, V. CRISPaint allows modular base-specific gene tagging using a ligase-4-dependent mechanism. *Nat Commun* **7**, 12338 (2016).
451. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res* **42**, e168 (2014).
452. Lindsay, H. et al. CrispRVariants charts the mutation spectrum of genome engineering experiments. *Nat Biotechnol* **34**, 701-702 (2016).
453. Gerstung, M. et al. Reliable detection of subclonal single-nucleotide variants in tumour cell populations. *Nat Commun* **3**, 811 (2012).
454. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235-242 (2011).
455. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010).
456. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501 (2010).
457. Ponder, J.W. & Case, D.A. Force fields for protein simulations. *Adv Protein Chem* **66**, 27-85 (2003).
458. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278 (2014).
459. Hing, Z.A. et al. Next-generation XPO1 inhibitor shows improved efficacy and in vivo tolerability in hematological malignancies. *Leukemia* **30**, 2364-2372 (2016).
460. Sorensen, M.M. & Henk, D. A first in class, first in human phase i trial of KPT-330 (selinexor), a selective inhibitor of nuclear export (SINE) in patients with advanced solid tumors. *ASCO Annual Meeting* (2014).
461. Vercruysse, T. et al. The second-generation exportin-1 inhibitor KPT-8602 demonstrates potent activity against acute lymphoblastic leukemia. *Clin Cancer Res* **23**, 2528-2541 (2017).
462. Etchin, J. et al. KPT-8602, a second-generation inhibitor of XPO1-mediated nuclear export, is well tolerated and highly active against AML blasts and leukemia-initiating cells. *Leukemia* **31**, 143-150 (2017).
463. Crochiere, M.L. et al. XPO1 target occupancy measurements confirm the selinexor recommended phase 2 dose. *Oncotarget* **8**, 110503-110516 (2017).
464. Singh, J., Petter, R.C., Baillie, T.A. & Whitty, A. The resurgence of covalent drugs. *Nat Rev Drug Discov* **10**, 307-317 (2011).
465. Zhou, S., Chan, E., Duan, W., Huang, M. & Chen, Y.Z. Drug bioactivation, covalent binding to target proteins and toxicity relevance. *Drug Metab Rev* **37**, 41-213 (2005).
466. Mosammaparast, N. & Pemberton, L.F. Karyopherins: from nuclear-transport mediators to nuclear-function regulators. *Trends Cell Biol* **14**, 547-556 (2004).
467. Forbes, D.J., Travesa, A., Nord, M.S. & Bernis, C. Nuclear transport factors: global regulation of mitosis. *Curr Opin Cell Biol* **35**, 78-90 (2015).
468. Wu, Z., Jiang, Q., Clarke, P.R. & Zhang, C. Phosphorylation of Crm1 by CDK1-cyclin-B promotes Ran-dependent mitotic spindle assembly. *J Cell Sci* **126**, 3417-3428 (2013).
469. Brodie, K.M. & Henderson, B.R. Characterization of BRCA1 protein targeting, dynamics, and function at the centrosome: a role for the nuclear export signal, CRM1, and Aurora A kinase. *J Biol Chem* **287**, 7701-7716 (2012).
470. Rousselet, A. Inhibiting Crm1 causes the formation of excess acentriolar spindle poles containing NuMA and B23, but does not affect centrosome numbers. *Biol Cell* **101**, 679-693 (2009).
471. Moriggi, G., Nieto, B. & Dosil, M. Rrp12 and the Exportin Crm1 participate in late assembly events in the nucleolus during 40S ribosomal subunit biogenesis. *PLoS Genet* **10**, e1004836 (2014).
472. Golomb, L. et al. Importin 7 and exportin 1 link c-Myc and p53 to regulation of ribosomal biogenesis. *Mol Cell* **45**, 222-232 (2012).
473. Bai, B., Moore, H.M. & Laiho, M. CRM1 and its ribosome export adaptor NMD3 localize to the nucleolus and affect rRNA synthesis. *Nucleus* **4**, 315-325 (2013).
474. Ernoult-Lange, M. et al. Nucleocytoplasmic traffic of CPEB1 and accumulation in Crm1 nucleolar bodies. *Mol Biol Cell* **20**, 176-187 (2009).

475. Tabe, Y. et al. Ribosomal biogenesis and translational flux inhibition by the selective inhibitor of nuclear export (SINE) XPO1 antagonist KPT-185. *PLoS One* **10**, e0137210 (2015).
476. Ines Tagoug, P.N., Jiri Slaby, Jacquelyn Babich, Justin Simms, Kathy Gratton, Li Ren, Nizar J. Bahlis XPO1 inhibition disrupts ribosomal subunits assembly and induces multiple myeloma (MM) cell death. *Blood* **122** (2013).
477. Maruyama, T. et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538-542 (2015).
478. Yu, C. et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* **16**, 142-147 (2015).
479. Oakes, B.L. et al. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat Biotechnol* **34**, 646-651 (2016).
480. van der Oost, J., Westra, E.R., Jackson, R.N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**, 479-492 (2014).
481. Maddalo, D. et al. In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* **516**, 423-427 (2014).
482. Han, K. et al. Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. *Nat Biotechnol* **35**, 463-474 (2017).
483. Najm, F.J. et al. Orthologous CRISPR-Cas9 enzymes for combinatorial genetic screens. *Nat Biotechnol* **36**, 179-189 (2018).
484. Kitzman, J.O., Starita, L.M., Lo, R.S., Fields, S. & Shendure, J. Massively parallel single-amino-acid mutagenesis. *Nat Methods* **12**, 203-206, 4 p following 206 (2015).
485. Soskine, M. & Tawfik, D.S. Mutational effects and the evolution of new protein functions. *Nat Rev Genet* **11**, 572-582 (2010).
486. Fowler, D.M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat Methods* **11**, 801-807 (2014).
487. Ma, L. et al. CRISPR-Cas9-mediated saturated mutagenesis screen predicts clinical drug resistance with improved accuracy. *Proc Natl Acad Sci U S A* **114**, 11751-11756 (2017).
488. Winters, I.P. et al. Multiplexed in vivo homology-directed repair and tumor barcoding enables parallel quantification of Kras variant oncogenicity. *Nat Commun* **8**, 2053 (2017).

Curriculum vitae

Jasper Edgar Neggers

Born in the Netherlands on June 12, 1989
+32 476 09 31 39
Pieter Coutereelstraat 76, Leuven
B3000, Belgium

neggersjasper@gmail.com

Education

2013-2018 **PhD training in Biomedical Sciences**
KU Leuven – Department of Microbiology and Immunology, Belgium
Group of Prof. Dr. Dirk Daelemans (Laboratory of Virology and Chemotherapy)

2011-2013 **Master of Science in Biomedical Sciences**
KU Leuven – Faculty of Medicine, Belgium
Group of Prof. Dr. Dirk Daelemans (Laboratory of Virology and Chemotherapy)
Thesis: "The Rev-mediated nuclear export of the unspliced HIV RNA"
Promotor: Prof. Dr. Dirk Daelemans

2008-2011 **Bachelor of Science in Biomedical Sciences**
KU Leuven – Faculty of Medicine, Belgium

2002-2008 **Voorbereidend Wetenschappelijk Onderwijs**
Stedelijk Gymnasium Breda – Breda, The Netherlands
Profiles: "Nature & Technology" and "Nature & Health"

Scientific publications

Neggers J.E., Kwanten B., Dierckx T., Noguchi H., Voet A., Bral L., Minner K., Massant B., Kint N., Delforge M., Vercruysse T., Baloglu E., Senapedis W., Jacquemyn M. and Daelemans D. Target identification of small molecules using large-scale CRISPR-Cas mutagenesis scanning of essential genes. *Nat. Commun., 9(1):502 (2018).*

Neggers J.E.*, Vercruysse T.*, De Bie J.*, Jacquemyn M., Vanstreels E., Schmid-Burgk J.L., Hornung V., Baloglu E., Landesman Y., Senapedis W., Shacham S., Dagklis A., Cools J. and Daelemans D. The second-generation exportin-1 inhibitor KPT-8602 demonstrates potent activity against acute lymphoblastic leukemia. *Clin. Cancer Res., 23(10):2528-2541 (2017).* *Equally contributing authors

Kim J., McMillan E., Kim H.S., Venkateswaran N., Makkar G., Rodriguez-Canales J., Villalobos P., **Neggers J.E.**, Mendiratta S., Wei S., Landesman Y., Senapedis W., Baloglu E., Chow C.B., Frink R.E., Gao B., Roth M., Minna J.D., Daelemans D., Wistuba I.I., Posner B.A., Scaglioni P.P. and White M.A. XPO1-dependent nuclear export is a druggable vulnerability in KRAS-mutant lung cancer. *Nature, 38(7623):114-117 (2016)*.

Neggers J.E., Vanstreels E., Baloglu E., Shacham S., Landesman Y. and Daelemans D. Heterozygous mutation of cysteine528 in XPO1 is sufficient for resistance to selective inhibitors of nuclear export. *Oncotarget, 7(42):68842-68850 (2016)*.

Boons E., Vanstreels E., Jacquemyn M., Nogueira T.C., **Neggers J.E.**, Vercruysse T., van den Oord J., Tamir S., Shacham S., Landesman Y., Snoeck R., Pannecouque C., Andrei G. and Daelemans D. Human exportin-1 is a target for combined therapy of HIV and AIDS related lymphoma. *EBioMedicine, 2(9):1102-1113 (2015)*.

Neggers J.E., Vercruysse T., Jacquemyn M., Vanstreels E., Baloglu E., Shacham S., Crochiere M., Landesman Y. and Daelemans D. Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing. *Chem. Biol., 22(1):107-116 (2015)*.

Conference proceedings

Neggers J.E., Kwanten B., Dierckx T., Noguchi H., Voet A., Bral L., Minner K., Massant B., Kint N., Delforge M., Vercruysse T., Baloglu E., Senapedis W., Jacquemyn M. and Daelemans D. Identification of a drug's molecular target using CRISPR/Cas-mediated mutagenesis scanning of essential proteins. *VIB - Genome Engineering and Synthetic Biology (3rd edition),* January 25-26, 2018. Poster Presentation.

Neggers J.E., Kwanten B., Dierckx T., Noguchi H., Voet A., Vercruysse T., Baloglu E., Senapedis W., Jacquemyn M. and Daelemans D. Identification of a drug's molecular target using CRISPR/Cas-mediated mutagenesis scanning of essential proteins. *Cell Symposium - CRISPR: From biology to Technology and Novel Therapeutics*, October 22-24, 2017. Selected for a short talk.

Neggers J.E., Kwanten B., Dierckx T., Noguchi H., Voet A., Vercruysse T., Baloglu E., Senapedis W., Jacquemyn M. and Daelemans D. Drug target identification using CRISPR-Cas-mediated directed evolution of essential proteins. *Horizon CRISPR Forum 2017*, October 17-18. Invited Speaker.

Neggers J.E., Kwanten B., Dierckx T., Vercruysse T., Jacquemyn M. and Daelemans D. Target identification using large-scale CRISPR/Cas-mediated mutagenesis scanning of essential genes. *Genome Engineering - The CRISPR/Cas revolution 2017 (Cold Spring Harbor Laboratory)*, July 21-24, 2017 (received travel stipend). Poster Presentation.

Neggers J.E., Kwanten B., Dierckx T. and Daelemans D. Identification of drug-target interactions through rapid selection of drug resistant protein variants generated by CRISPR/Cas9-induced NHEJ repair. *KU Leuven Oncoforum*, Jun 9, 2017. Poster presentation (won poster prize).

Neggers J.E., Kwanten B., Dierckx T. and Daelemans D. Identification of drug-target interactions through rapid selection of drug resistant protein variants generated by CRISPR/Cas9-induced NHEJ Repair. *American Association for Cancer Research Annual Meeting 2017*, Apr 5, 2017. Poster presentation.

De Bie J., **Neggers J.E.**, Dagklis A., Vercruysse T., Gielen O., Baloglu E., Senapedis W., Shacham S., Landesman Y., Uyttebroeck A., Boeckx N., Cools J. and Daelemans D. KPT-8602 is a second-generation XPO1 inhibitor with improved in vivo tolerability and potent in vivo activity against acute lymphoblastic leukemia. *21st Congress of the European Hematology Association*, Jun 8, 2016. Poster presentation.

Daelemans D., **Neggers J.E.,** De Bie J., Jacquemyn M., D'Hoore A., Vanstreels E., Baloglu E., Landesman Y., Shacham S., Senapedis W., Dagklis A., Vercruysse T. and Cools J. KPT-8602 is a second-generation XPO1 inhibitor with improved in vivo tolerability that demonstrates potent acute lymphoblastic leukemia activity. *American Association for Cancer Research Annual Meeting 2016*, Apr 17, 2016. Poster presentation.

Daelemans D., Boons E., Vanstreels E., Jacquemyn M., Nogueira T.C., **Neggers J.E.**, Vercruysse T., van den Oord J., Tamir S., Kauffman M., Shacham S., Landesman Y., Snoeck R., Pannecouque C. and Andrei G. Selective Inhibitors of Nuclear Export (SINE) Compounds Suppress Both HIV Replication and AIDS Related Lymphoma. *57th American Society of Hematology Annual Meeting*, Dec 6, 2015. Poster presentation.

Neggers J.E., Landesman Y., Shacham S. and Daelemans D. A single amino acid substitution by CRISPR/Cas9 genome editing at cysteine 528 of XPO1 confers resistance of malignant cells to selective inhibitors of nuclear export. *Genome Engineering - The CRISPR/Cas revolution (Cold Spring Harbor Laboratory)*, Sep 26, 2015 (received travel stipend). Poster presentation.

Neggers J.E., Vercruysse T., Jacquemyn M., Vanstreels E., Baloglu E., Shacham S., Crochiere M., Landesman Y., Daelemans D. CRISPR/Cas9 genome editing validates CRM1/XPO1 as target for clinical inhibitors of nuclear export. *KU Leuven Oncoforum*, Jun 5, 2015. Poster presentation.

Neggers J.E., Vercruysse T., Jacquemyn M., Vanstreels E., Baloglu E., Shacham S., Crochiere M., Senapedis W., Landesman Y. and Daelemans D. XPO1 is selinexor's prime target: validation by mutating cysteine 528 on both XPO1 alleles using CRISPR/Cas9 genome editing. *American Association for Cancer Research Annual Meeting 2015*, Apr 20, 2015. Poster presentation.